English Translation
of Certified Copy

JAPAN PATENT OFFICE

This is to certify that the annexed is a true copy of the following application as filed with this Office.

Date of Application : November 7, 2003

Application Number : Patent Application No. 2003-378039

Applicant : SANKYO COMPANY, LIMITED

Date : September 24, 2004

Commissioner, Hiroshi Ogawa
Patent Office

Official Seal

Certificate Serial No.

Name of Document	Patent Application
Docket Number	2003175SU
Address To :	Commissioner, Patent Office
International Patent Classification	C12Q1/68

Inventor
- Address or Domicile: c/o SANKYO COMPANY, LIMITED
  2-58, Hiromachi 1-chome, Shinagawa-ku, Tokyo, Japan
- Name: Makoto Koizumi

Patent Applicant
- Identification Number: 000001856
- Name: SANKYO COMPANY, LIMITED

Agent
- Identification Number: 100081400
- Patent Attorney
- Name: Akio Ohno

Appointed Agent
- Identification Number: 100092716
- Patent Attorney
- Name: Yasuo Nakada

Appointed Agent
- Identification Number: 100115750
- Patent Attorney
- Name: Toshiaki Yaguchi

Appointed Agent
- Identification Number: 100119622
- Patent Attorney
- Name: Reiko Kinpara

Appointed Agent
- Identification Number: 100125025
- Patent Attorney
- Name: Tomoki Echigo

Indication of Official Fee

Prepayment Register Number	010216
Amount to be paid	21,000 yen

List of Materials to be submitted

Name of Material	Claims for the Patent	1
Name of Material	Specification	1
Name of Material	Drawings	1
Name of Material	Abstract	1
Number of General Power of Attorney		9704937
Number of General Power of Attorney		9704935
Number of General Power of Attorney		0113519
Number of General Power of Attorney		0113520
Number of General Power of Attorney		0302858

[Title of the Document] Claims for the Patent

[Claim 1]

An oligonucleotide,

(a) wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, and the other nucleotides are natural nucleotides; and

(b) wherein the oligonucleotide has a nucleotide complementary to the reference nucleotide of a target gene at the 3'-end position thereof, and has nucleotides complementary to the nucleotide sequence of the target gene at the other positions,

or a salt thereof.

[Claim 2]

An oligonucleotide,

(a) wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, and the other nucleotides are natural nucleotides; and

(b) wherein the oligonucleotide has a nucleotide complementary to the mutant nucleotide of a target gene at the 3'-end position thereof, and has nucleotides complementary to the nucleotide sequence of the target gene at the other positions,

or a salt thereof.

[Claim 3]

An oligonucleotide according to claim 1 or 2, characterized in that the polymorphism is monobasic polymorphism, or a salt thereof.

[Claim 4]

An oligonucleotide or a salt thereof according to any one of claims 1 to 3, characterized by having a base length of 18 to 25 bases.

[Claim 5]

A method for detecting gene polymorphism, characterized by using an oligonucleotide according to any one of claims 1 to 4.

[Claim 6]

A method for determining the nucleotide sequence of a genetically polymorphic sequence, characterized by using an oligonucleotide according to any one of claims 1 to 4.

[Claim 7]

A method for detecting gene polymorphism, comprising the following steps (a) and (b):

(a) a step of performing PCR with nucleic acid comprising a genetically polymorphic sequence as a template using an oligonucleotide according to any one of claims 1 to 4 and an oligonucleotide capable of amplifying a sequence of interest together with said oligonucleotide in the PCR; and

(b) a step of determining the presence or absence of gene polymorphism in the nucleic acid based on whether or not a reaction product can be generated in step (a).

[Claim 8]

A method for determining the nucleotide sequence of a genetically polymorphic sequence, comprising the following steps (a) and (b):

(a) a step of performing PCR with nucleic acid comprising a genetically polymorphic sequence as a template using an

oligonucleotide according to any one of claims 1 to 4 and an oligonucleotide capable of amplifying the sequence of interest together with said oligonucleotide in the PCR; and

(b) a step of determining the nucleotide sequence of a genetically polymorphic sequence in the nucleic acid based on whether or not a reaction product can be generated in step (a).

[Claim 9]

A method according to claim 7 or 8, characterized in that the detection of the presence or absence of generation of a reaction product uses one or more method selected from the group consisting of electrophoresis, Taq-Man PCR, and a MALDI-TOF/MS method.

[Claim 10]

A method according to any one of claims 5 to 9, characterized in that the gene polymorphism is a single nucleotide polymorphism.

[Claim 11]

A kit for detecting gene polymorphism, which comprises the following (a) to (d):

(a) an oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other nucleotides are natural nucleotides, the 3'-end nucleotide thereof is a nucleotide complementary to the reference nucleotide of a target gene, and at the other positions the nucleotides are complementary to the nucleotide sequence of the target gene;

(b) an oligonucleotide capable of amplifying a sequence of interest, together with the oligonucleotide described in (a) above;

(c) DNA polymerase; and

(d) a PCR buffer.

[Claim 12]

A kit for detecting gene polymorphism, comprising the following (a) to (d):

(a) an oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other nucleotides are natural nucleotides, the 3'-end nucleotide thereof is a nucleotide complementary to the mutant nucleotide of a target gene, and at the other positions the nucleotides are complementary to the nucleotide sequence of the target gene;

(b) a primer capable of amplifying a sequence of interest, together with the oligonucleotide described in (a) above;

(c) DNA polymerase; and

(d) a PCR buffer.

[Claim 13]

A kit for detecting gene polymorphism, comprising the following (a) to (e):

(a) an oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other nucleotides are natural nucleotides, the 3'-end nucleotide thereof is a nucleotide complementary to the reference nucleotide of a target gene, and at the other

positions the nucleotides are complementary to the nucleotide sequence of the target gene;

(b) an oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other nucleotides are natural nucleotides, the 3'-end nucleotide thereof is a nucleotide complementary to the mutant nucleotide of a target gene, and at the other positions the nucleotides are complementary to the nucleotide sequence of the target gene;

(c) an oligonucleotide capable of amplifying a sequence of interest, together with the oligonucleotide described in (a) or (b) above;

(d) DNA polymerase; and

(e) a PCR buffer.

[Claim 14]

A kit for detecting gene polymorphism according to any one of claims 11 to 13, characterized in that the oligonucleotide, and the oligonucleotide capable of amplifying a sequence of interest together with said oligonucleotide, each have a base length of 18 to 25 bases.

[Claim 15]

A kit according to any one of claims 11 to 14, characterized in that the gene polymorphism is a single nucleotide polymorphism.

[Title of the Document] Specification

[Title of the Invention] METHOD OF DETECTING GENETIC POLYMORPHISM

[Technical Field]

[0001]

The present invention relates to a method for detecting gene polymorphism by PCR using an oligonucleotide comprising an ENA unit, an oligonucleotide used in detection of gene polymorphism, and a kit for detecting gene polymorphism comprising the above oligonucleotide.

[Background Art]

[0002]

As a result of advances in pharmacogenomics, it is now possible to predict the effects or side effects of drugs in individual patients by gene diagnosis based on the relationship between gene polymorphism and drug effects, or between gene polymorphism and side effects. An example is the gene polymorphism of drug metabolizing enzymes. Examples of known drug metabolizing enzymes whose activity is increased or decreased by such polymorphism include cytochrome P4501A2, cytochrome P4502A6, cytochrome P4502C9, cytochrome P4502C19, cytochrome P4502D6, and cytochrome P4502E1. In addition, it has been reported that among a group of enzymes known as conjugation enzymes, such as thiopurine methyltransferase, N-acetyltransferase, UDP-glucuronosyltransferase, or glutathione S-transferase, such gene polymorphism exists, and that the activities of the above enzymes are decreased by such polymorphism (see Non-Patent Document 1, for example).

[0003]

Moreover, by examining the relationship between gene polymorphism and diseases, the pre-diagnosis of several diseases or the determination of prognosis becomes possible. A large number of disease-associated genes discovered as a result of polymorphism analyses have been reported. Examples of such disease-associated genes, which have been reported, include: HLA, a causative gene of ulcerative colitis; TCRα, a causative gene of rheumatoid arthritis; APOE4, a causative gene of Alzheimer's disease; a dopamine D3 receptor, a causative gene of schizophrenia; tryptophan hydroxylase, a causative gene of manic-depressive psychosis; an angiotensin precursor, a causative gene of albuminuria; blood coagulation factor VII, a causative gene of myocardial infarct; and leptin, a causative gene of adiposis (see Non-Patent Document 2, for example).

[0004]

Examples of methods for detecting gene polymorphism, which have been developed, include: the PCR-RFLP method, involving a combination of the polymerase chain reaction (PCR) method and cleavage with restriction enzymes (see Non-Patent Document 3, for example); the SSCP (single-strand conformation polymorphism) method, based on the principle that single-strand DNA and RNA having different sequences exhibit different electrophoretic mobility in polyacrylamide gels; and the AS-PCR (allele-specific PCR) method, based on the principle that mismatches existing at the 3'-end of an oligonucleotide primer inhibit elongation of the primer.

[0005]

Since the PCR-RFLP method comprises a treatment with restriction enzymes for 3 to 24 hours in its test process, it is difficult to say that this is a rapid method. The SSCP method is excellent in that when one or several mutations exist in any part of the nucleotide sequence used as a test target, this method is able to detect such existence at high sensitivity. However, since the experimental conditions are strictly controlled to detect a subtle difference in mobility, this is extremely complicated, and furthermore, the position of the mutation cannot be identified by this method. In addition, in order to perform the SSCP method using actual analytes such as blood or tissues, it is necessary to prepare, in advance, a large amount of nucleic acid via cloning or the PCR method. Thus, this method is not suitable for efficiently testing a large number of analytes.

[0006]

The AS-PCR method is a method that involves a modification of PCR. For this method it is not necessary to prepare in advance a large amount of nucleic acid. This method is based on the fact that an amplified product can be obtained only when primers having no mismatch around the 3'-ends thereof are used. This is a method suitable for efficiently testing a large number of analytes. However, there are cases where such an amplified product can be obtained in ordinary PCR, even when mismatches exist in primers. Thus, the above method has been problematic in terms of stringency.

[0007]

As a result of studies directed towards solving the aforementioned problems with polymorphism detection methods, the present inventors have found that when an oligonucleotide used as a PCR primer has a polymorphic portion at the third position from the 3'-end, which has been modified with ENA, the amount of an amplified product generated due to mismatches is decreased, and gene polymorphism can be detected with high precision. In addition, the inventors have further provided a kit for use in the above detection method, thereby completing the present invention.

[Non-Patent Document 1] "SNP *Idenshi Takei no Senryaku* (Strategy of SNP Gene Polymorphism)", edited by Yusuke Nakamura, Nakayama Shoten, June 5, 2000

[Non-Patent Document 2] "Nature Genetics", 1999, Vol. 22, pp.139-144

[Non-Patent Document 3] "Science", 1991, Vol. 252, pp.16-43

[Disclosure of the Invention]

[Problems to be Solved by the Invention]

[0008]

It is an object of the present invention to provide a method for detecting gene polymorphism, an oligonucleotide for use in the above method, and a gene polymorphism detection kit comprising the above oligonucleotide.

[Means for Solving the Problems]

[0009]

The present invention relates to a method for detecting gene polymorphism, which utilizes a phenomenon whereby during

the synthesis of a synthetic oligonucleotide primer having a nucleotide sequence complementary to the nucleotide sequence of nucleic acid used as a template, if the nucleotide at the 3'-end of the synthetic oligonucleotide primer is a nucleotide that is not complementary to the nucleotides of the template, elongation of the primer with DNA polymerase does not take place, but if a synthetic oligonucleotide primer is used that is completely complementary to the nucleotide sequence of the nucleic acid used as a template, elongation of the primer with DNA polymerase takes place.

[0010]

More specifically, the present invention relates to a method for detecting gene polymorphism, characterized by using, as a primer, a synthetic oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide (ENA) unit.

The present invention specifically includes the following features:

(1) An oligonucleotide,

(a) wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, and the other nucleotides are natural nucleotides; and

(b) wherein the oligonucleotide has a nucleotide complementary to the reference nucleotide of a target gene at the 3'-end position thereof, and has nucleotides complementary to the nucleotide sequence of the target gene at the other positions,

or a salt thereof.

(2) An oligonucleotide,

(a) wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, and the other nucleotides are natural nucleotides; and

(b) wherein the oligonucleotide has a nucleotide complementary to the mutant nucleotide of a target gene at the 3'-end position thereof, and has nucleotides complementary to the nucleotide sequence of the target gene at the other positions,

or a salt thereof.

(3) An oligonucleotide according to (1) or (2) above, characterized in that the polymorphism is monobasic polymorphism, or a salt thereof.

(4) An oligonucleotide or a salt thereof according to any one of (1) to (3) above, characterized by having a base length of 18 to 25 bases.

(5) A method for detecting gene polymorphism, characterized by using an oligonucleotide according to any one of (1) to (4) above.

(6) A method for determining the nucleotide sequence of a genetically polymorphic sequence, characterized by using an oligonucleotide according to any one of (1) to (4) above.

(7) A method for detecting gene polymorphism, comprising the following steps (a) and (b):

(a) a step of performing PCR with a nucleic acid comprising a genetically polymorphic sequence as a template using an oligonucleotide according to any one of (1) to (4) above and an oligonucleotide capable of amplifying a sequence of

interest together with the aforementioned oligonucleotide in the PCR; and

(b) a step of determining the presence or absence of gene polymorphism in the nucleic acid based on whether or not a reaction product can be generated in step (a).

(8) A method for determining the nucleotide sequence of a genetically polymorphic sequence, comprising the following steps (a) and (b):

(a) a step of performing PCR with a nucleic acid comprising a genetically polymorphic sequence as a template using an oligonucleotide according to any one of (1) to (4) above and an oligonucleotide capable of amplifying a sequence of interest together with the aforementioned oligonucleotide in the PCR; and

(b) a step of determining the nucleotide sequence of a genetically polymorphic sequence in the nucleic acid based on whether or not a reaction product can be generated in step (a).

(9) A method according to (7) or (8), characterized in that the detection of the presence or absence of generation of a reaction product uses one or more method selected from the group consisting of electrophoresis, Taq-Man PCR, and a MALDI-TOF/MS method.

(10) A method according to any one of (5) to (9) above, characterized in that the gene polymorphism is a single nucleotide polymorphism.

(11) A kit for detecting gene polymorphism, which comprises the following (a) to (d):

(a) an oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other nucleotides are natural nucleotides, the 3'-end nucleotide thereof is a nucleotide complementary to the reference nucleotide of a target gene, and at the other positions the nucleotides are complementary to the nucleotide sequence of the target gene;

(b) an oligonucleotide capable of amplifying a sequence of interest, together with the oligonucleotide described in (a) above;

(c) DNA polymerase; and

(d) a PCR buffer.

(12) A kit for detecting gene polymorphism, comprising the following (a) to (d):

(a) an oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other nucleotides are natural nucleotides, the 3'-end nucleotide thereof is a nucleotide complementary to the mutant nucleotide of a target gene, and at the other positions the nucleotides are complementary to the nucleotide sequence of the target gene;

(b) a primer capable of amplifying a portion of a sequence of interest, together with the oligonucleotide described in (a) above;

(c) DNA polymerase; and

(d) a PCR buffer.

(13) A kit for detecting gene polymorphism, comprising the following (a) to (e):

(a) an oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other nucleotides are natural nucleotides, the 3'-end nucleotide thereof is a nucleotide complementary to the reference nucleotide of a target gene, and at the other positions the nucleotides are complementary to the nucleotide sequence of the target gene;

(b) an oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other nucleotides are natural nucleotides, the 3'-end nucleotide thereof is a nucleotide complementary to the mutant nucleotide of a target gene, and at the other positions the nucleotides are complementary to the nucleotide sequence of the target gene;

(c) an oligonucleotide capable of amplifying a sequence of interest, together with the oligonucleotide described in (a) or (b) above;

(d) DNA polymerase; and

(e) a PCR buffer.

(14) A kit for detecting gene polymorphism according to any one of (11) to (13) above, characterized in that the oligonucleotide, and the oligonucleotide capable of amplifying a sequence of interest together with the aforementioned oligonucleotide, each have a base length of 18 to 25 bases.

(15) A kit according to any one of (11) to (14) above, characterized in that the gene polymorphism is a single nucleotide polymorphism.

The principle of the method for detecting gene polymorphism of the present invention is as follows:

[0011]

The 3'-end of a primer is directed to the polymorphic portion of a sequence in which gene polymorphism is intended to be detected (sequence of interest), and the third nucleotide from the 3'-end of the primer is modified with an ENA unit. When this primer and nucleic acid containing a nucleotide sequence which is the gene polymorphism detection target are reacted with a mixture of nucleic acid synthesizing enzymes in a reaction solution, if the 3'-end of the primer matches (i.e. the base is complementary), a nucleic acid synthesis reaction takes place. In contrast, if the 3'-end does not match, a nucleic acid synthesis reaction does not take place. Utilizing the difference between the situation in which a nucleic acid synthesis reaction takes place when the 3'-end matches and the situation when a reaction does not take place when the 3'-end does not match, mutation can be detected in a nucleotide sequence. This principle is explained in Figures 1 and 2.

[0012]

Figure 1 shows the situation when there is no mutation (polymorphism) in a nucleic acid sequence. (i) is a template nucleic acid of a target in which it is intended to examine mutation (polymorphism) in the nucleic acid sequence, and it has the sequence ATGC as a nucleotide sequence portion thereof. This template nucleic acid is annealed with oligonucleotide (ii) in which the third position from the

3'-end thereof has been modified with ENA (a 2'-*O*,4'-*C*-ethylene-5-methyluridine unit is represented by eT), so as to form a double strand. In this case, at least the 3'-end of the nucleotide sequence of (ii) has a structure that is complementary to the corresponding base, and (ii) and (i) form a complementary strand. The 3'-end portion of oligonucleotide forming such a complementary strand is recognized by nucleic acid synthesizing enzyme (iii), and the nucleic acid synthesis reaction is continued. Specific nucleotide sequences shown in the figure are used for explanation, and thus, it does not mean that the present invention is effective only for such nucleotide sequences.
[0013]

Figure 2 shows a case where there is mutation (polymorphism) in a nucleic acid sequence. (i) is a template nucleic acid of a target in which it is intended to examine mutation (polymorphism) in the nucleic acid sequence, and it has the sequence ATAC as a nucleotide sequence portion thereof. This template nucleic acid is annealed with oligonucleotide (ii) in which the third position from the 3'-end thereof has been modified with ENA (a 2'-*O*,4'-*C*-ethylene-5-methyluridine unit is represented by eT), so as to form a double strand. In this case, at least the 3'-end of the nucleotide sequence of (ii) does not have a structure that is complementary to the corresponding base, and C at the 3'-end of (ii) does not form a complementary strand. The 3'-end portion of oligonucleotide that does not form a complementary strand cannot be recognized by the nucleic

acid synthesizing enzyme (iii), and the nucleic acid synthesis reaction cannot be continued. Specific nucleotide sequences shown in the figure are used for explanation, and thus, it does not mean that the present invention is effective for only for such nucleotide sequences.

[Advantages of the Invention]

[0014]

The present invention provides a method for detecting gene polymorphism with higher precision than that in the case of using natural oligonucleotides. Moreover, the present invention also provides an oligonucleotide for use in detection of gene polymorphism and a kit for detecting gene polymorphism, which comprises the above oligonucleotide, which can be used for the above method.

[Best Mode for Carrying Out the Invention]

[0015]

1. Explanation of terms

The term "gene polymorphism" is used in the present specification to mean a certain gene locus, which comprises (a) substitution of a single base with another base (single nucleotide polymorphism (SNP)) and/or (b) deletion or insertion of one to several tens of bases (wherein the number of bases is several thousands of bases in some cases) (insertion/deletion polymorphism). In the present specification, single nucleotide polymorphism is also referred to as SNP, and it means a difference of a single base in nucleotide sequences between individuals.

[0016]

It is known that alternative nucleotides may be present at a single nucleotide polymorphism position (for example adenine or guanine, thymine or cytosine, etc). The ratio of these variants differs depending on the target gene. The term "target gene" is used in the present specification to mean a gene used as a target for gene polymorphism.

[0017]

In the present specification, a sequence containing a nucleotide with a high frequency of occurrence among alternative variants in a single nucleotide polymorphic site of a target gene is defined as a reference sequence, and the nucleotide in the single nucleotide polymorphic site in such a reference sequence is defined as a reference nucleotide. On the other hand, a sequence containing a nucleotide with a low frequency of occurrence is defined as a mutant sequence, and the nucleotide in the single nucleotide polymorphic site in such a mutant sequence is defined as a mutant nucleotide.

[0018]

Moreover, when the polymorphism is a deletion polymorphism, a sequence having no deletion is defined as the reference sequence, and a sequence having a deletion is defined as the mutant sequence.

[0019]

Furthermore, when the polymorphism is an insertion polymorphism, a sequence having no insertion is defined as the reference sequence, and a sequence having an insertion is defined as the mutant sequence.

[0020]

In the present specification, the expression "have polymorphism" is used to mean that a sequence comprising a polymorphism of interest in a target gene has a mutant sequence, and the expression "does not have polymorphism" is used to mean that a sequence comprising a polymorphism of interest in a target gene is a reference sequence.

In the present specification, the term "natural nucleotide" includes adenine nucleotide, guanine nucleotide, cytosine nucleotide, uracil nucleotide, and thymine nucleotide. In addition, the term "natural oligonucleotide" is used to mean an oligonucleotide composed of natural nucleotides such as adenine nucleotide, guanine nucleotide, cytosine nucleotide, uracil nucleotide, or thymine nucleotide.

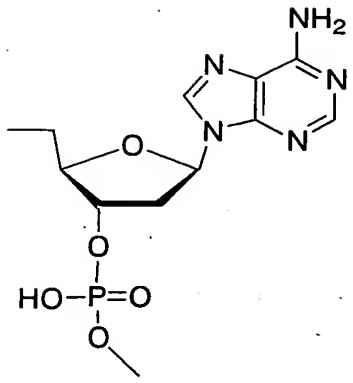
[0021]

In the present specification, adenine nucleotide may be represented by A^p, guanine nucleotide is represented by G^p, cytosine nucleotide is represented by C^p, and thymine nucleotide is represented by T^p. Moreover, with regard to the nucleotide at the 3'-end of a natural oligonucleotide, adenine nucleoside is represented by A^t, guanine nucleoside is represented by G^t, cytosine nucleoside is represented by C^t, and thymine nucleoside may be represented by T^t.

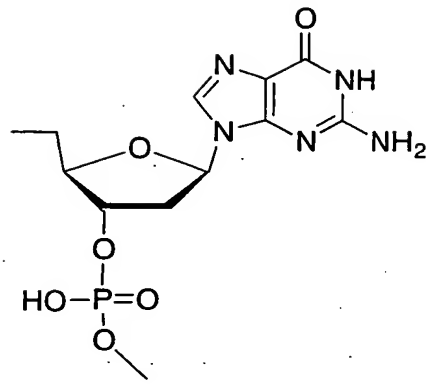
The structural formulas of natural nucleotides are shown below.

[0022]

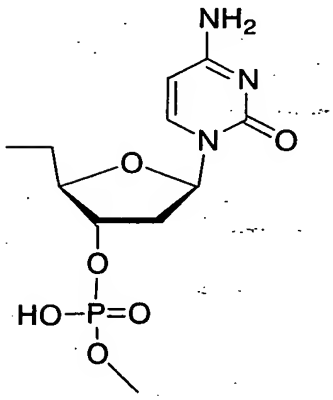
[Chemical formula 1]



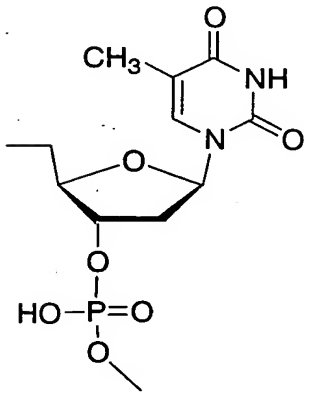
A^p



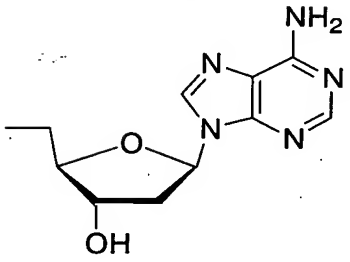
G^p



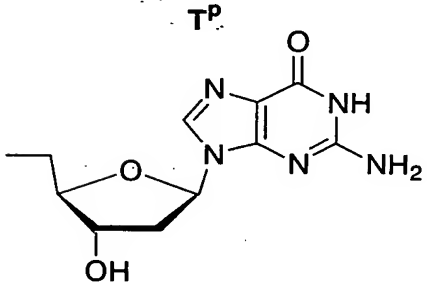
C^p



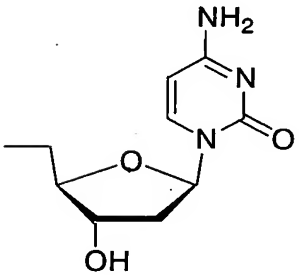
T^p



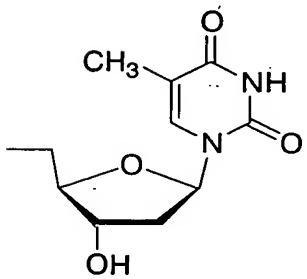
A^t



G^t



C^t



T^t

[0023]

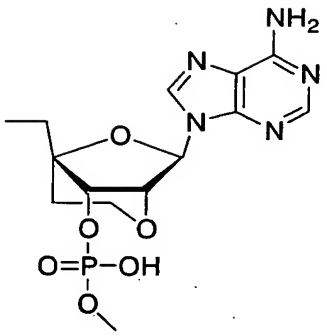
In the present specification, the term "ENA nucleotide" (hereinafter also referred to as "ENA") is used to mean a nucleotide, which is formed by crosslinking the oxygen atom at the 2' position of a sugar portion and the carbon atom at the 4' position thereof using an ethylene chain (refer to Japanese Patent No. 3420984). An oligonucleotide, into which ENA has been introduced, has high binding ability to complementary strand RNA. In addition, the ENA nucleotides are characterized in that they have a higher resistance to nuclease than LNA nucleotides (2'-*O*,4'-*C*-methylene nucleotide (Japanese Patent Laid-Open No. 10-304889), which are formed by crosslinking, with a methylene chain, an oxygen atom at the 2' position and a carbon atom at the 4' position of a sugar portion (Bioorg. Med. Chem. Lett. 12, 73-76, (2002)).

[0024]

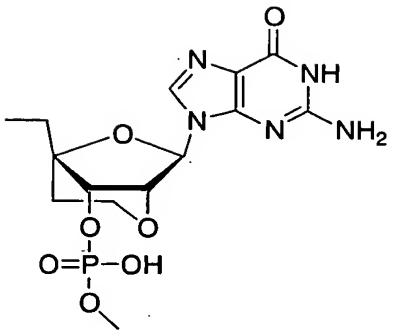
In the present specification, the term "ENA unit" is used to mean A^{e2p}, G^{e2p}, C^{e2p}, $5C^{e2p}$ or T^{e2p}, or when the 3'-end of an oligonucleotide has such a unit, or when ENA is treated as a nucleoside, it means any group selected from among C^{e2t}, $5C^{e2t}$, and T^{e2t}. The structures thereof are shown below. Further, the structure of C^{e1p} is also shown as an LNA unit below.

[0025]

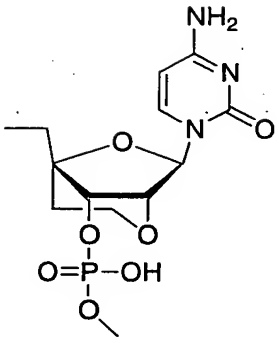
[Chemical formula 2]



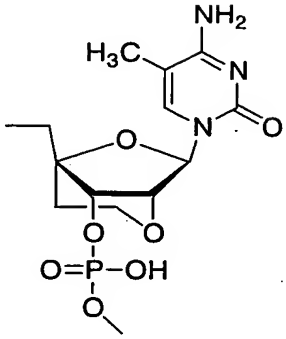
A^{e2p}



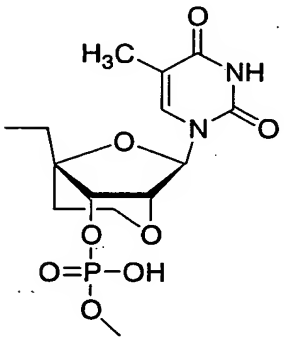
G^{e2p}



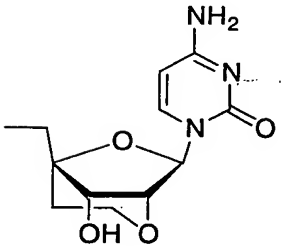
C^{e2p}



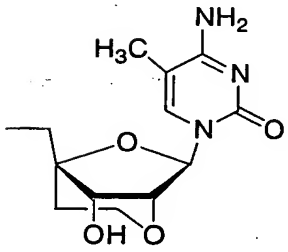
$5C^{e2p}$



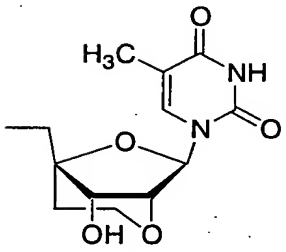
T^{e2p}



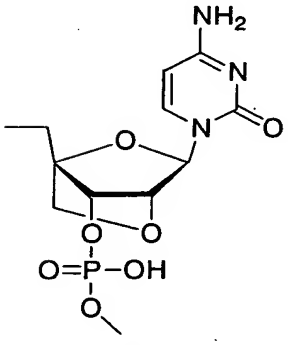
C^{e2t}



$5C^{e2t}$



T^{e2t}



C^{e1p}

[0026]

The term "salt thereof" is used in the present specification to mean a salt, into which a compound of the present invention can be converted. Preferred examples of such a salt include: alkali metal salts such as a sodium salt, a potassium salt, or a lithium salt; alkali-earth metal salts such as a calcium salt or a magnesium salt; metal salts such as an aluminum salt, an iron salt, a zinc salt, a copper salt, a nickel salt, or a cobalt salt; inorganic salts such as an ammonium salt; amine salts including organic salts such as t-octylamine salt, a dibenzylamine salt, a morpholine salt, a glucosamine salt, a phenylglycine alkyl ester salt, an ethylenediamine salt, an N-methylglucamine salt, a guanidine salt, a diethylamine salt, a triethylamine salt, a dicyclohexylamine salt, an N,N'-dibenzylethylenediamine salt, a chloroprocaine salt, a procaine salt, a diethanolamine salt, an N-benzyl-phenethylamine salt, a piperazine salt, a tetramethylammonium salt, or a tris(hydroxymethyl)aminomethane salt; halogenated hydroacid salts such as hydrofluoride, hydrochloride, hydrobromate, or hydroiodide; inorganic acid salts such as nitrate, perchlorate, sulfate, or phosphate; lower alkanesulfonates such as methanesulfonate, trifluoromethanesulfonate, or ethanesulfonate; arylsulfonates such as benzenesulfonate or p-toluenesulfonate; organic acid salts such as acetate, malate, fumarate, succinate, citrate, tartrate, oxalate, or

maleate; and amino acid salts such as a glycine salt, a lysine salt, an arginine salt, an ornithine salt, a glutamate salt, or aspartate salt.

[0027]

A compound of the present invention and a salt thereof can also be present in the form of a hydrate. The present invention includes such hydrates.

[0028]

2. Analyte

As an analyte used as a target wherein gene polymorphism is to be detected in the present invention, a sample containing nucleic acid can be used. An example of such nucleic acid is genomic DNA.

[0029]

For example, in order to detect human gene polymorphism, an analyte containing human genomic DNA can be used. In order to detect mouse gene polymorphism, mouse genomic DNA can be used. Hereafter, human genomic DNA will be used as an example. However, genomic DNA derived from other organisms can also be obtained in the same manner.

[0030]

As material for obtaining genomic DNA, all types of cells (excluding germ cells), tissues, organs, etc. collected from a subject, can be used. Preferred examples of such materials include leukocytes or monocytes separated from peripheral blood, the most preferred example being leukocytes. Such materials can be collected by methods commonly used in clinical tests.

[0031]

When leukocytes are used, the leukocytes are first separated from the peripheral blood collected from a subject by well known methods. Subsequently, proteinase K and sodium dodecyl sulfate (SDS) are added to the leukocytes so obtained to digest and denature proteins, this is followed by phenol/chloroform extraction to obtain genomic DNA (including RNA). RNA is removed with RNase, as necessary. However, the present invention is not limited to the aforementioned method. In order to extract genomic DNA from a sample containing human genomic DNA, methods publicly known in the technical field of the present invention are preferred, such as methods described in publications (refer to Sambrook, J. et al. (1989): "Molecular Cloning: A Laboratory Manual (2nd Ed.)" Cold Spring Harbor Laboratory, NY, for example), or a method using a commercially available DNA extraction kit.

[0032]

The purity of an analyte containing DNA is not particularly limited, as long as it can be used for PCR. A crude extract, a purified product, etc. obtained from a sample can be used.

[0033]

3. Selection of target gene

Any gene can be used as a target for detection of wherein gene polymorphism, provided that at least a nucleotide sequence portion thereof is already known and polymorphism exists in that portion. Examples of such target genes include known drug metabolizing genes associated with drug effects,

or the side effects of drugs, such as cytochrome P4501A2, cytochrome P4502A6, cytochrome P4502C9, cytochrome P4502C19, cytochrome P4502D6, or cytochrome P4502E1. In addition, further examples include thiopurine methyltransferase, N-acetyltransferase, UDP-glucuronosyltransferase, glutathione S-transferase, and disease-associated genes such as HLA that is a causative gene of ulcerative colitis, TCRα that is a causative gene of rheumatoid arthritis, APOE4 that is a causative gene of Alzheimer's disease, a dopamine D3 receptor that is a causative gene of schizophrenia, tryptophan hydroxylase that is a causative gene of manic-depressive psychosis, an angiotensin precursor that is a causative gene of albuminuria, blood coagulation factor VII that is a causative gene of myocardial infarct, or leptin that is a causative gene of adiposis. A further example is human prothrombin.

[0034]

Examples of polymorphism include deletion polymorphism on polymorphism of a mouse angiopoietin-like protein 3 promoter.

[0035]

The position of the polymorphism in a gene may be any of: a translated region, a nontranslated region, a regulatory region such as a promoter, or an intron, and other regions.

[0036]

4. Oligonucleotide primer

The following oligonucleotides can be synthesized using an automated nucleic acid synthesizer.

[0037]

Natural oligonucleotides can be synthesized using natural phosphoramidite. A 2'-*O*,4'-*C*-ethylene nucleotide can be synthesized using the following: (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-6-N-benzoyladenosine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite (Example 14 of Japanese Patent No. 3420984), 5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-2-N-isobutyrylguanosine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite (Example 27), 5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-4-N-benzoylcytidine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite (Example 5), 5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-4-N-benzoyl-5-methylcytidine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite (Example 22), and 5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-5-methyluridine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite (Example 9).

[0038]

(1) Oligonucleotides used in detection of gene polymorphism

Oligonucleotides used in detection of gene polymorphism in the present invention include the following (a) and/or (b):

[0039]

(a) an oligonucleotide having a nucleotide sequence complementary to a reference sequence,

(i) wherein the third nucleotide from the 3'-end thereof (when the nucleotide at the 3'-end is defined as the first nucleotide) is a 2'-*O*,4'-*C*-ethylene nucleotide (ENA) unit, and the other nucleotides are natural nucleotides;
(ii) wherein the oligonucleotide has a nucleotide complementary to the reference nucleotide of a target gene at the 3'-end position thereof, and has nucleotides complementary to the nucleotide sequence of the target gene at the other positions; and
(iii) wherein the length of the oligonucleotide primer is not particularly limited, as long as it is able to amplify nucleic acid in PCR, but the length is preferably 15 to 40 nucleotides, more preferably 18 to 35 nucleotides, and still more preferably 18 to 25 nucleotides.

[0040]

(b) an oligonucleotide having a nucleotide sequence complementary to a mutant sequence,
(i) wherein the third nucleotide from the 3'-end thereof (when the nucleotide at the 3'-end is defined as the first nucleotide) is a 2'-*O*,4'-*C*-ethylene nucleotide (ENA) unit, and the other nucleotides are natural nucleotides;
(ii) wherein the oligonucleotide has a nucleotide complementary to the mutant nucleotide of a target gene at the 3'-end position thereof, and has nucleotides complementary to the nucleotide sequence of the target gene at the other positions; and
(iii) wherein the length of the oligonucleotide thereof is not particularly limited, as long as it is able to amplify

a nucleic acid in PCR, but the length is preferably 15 to 40 nucleotides, more preferably 18 to 35 nucleotides, and still more preferably 18 to 25 nucleotides.

[0041]

The fact that the third nucleotide from the 3'-end of the oligonucleotide is a 2'-*O*,4'-*C*-ethylene nucleotide (ENA) unit in (a) and (b) above means that the third nucleotide is not a natural nucleotide but an ENA nucleotide. It means, for example, that A^{e2p} is used instead of A^{p}, G^{e2p} is used instead of G^{p}, $5C^{e2p}$ or C^{e2p} is used instead of C^{p}, and T^{e2p} is used instead of T^{p}, for example.

[0042]

These oligonucleotides may be referred to as forward primers.

[0043]

(2) Oligonucleotides used as a pair

(a) Oligonucleotides used in PCR

The sequence of an oligonucleotide used in PCR as a part of a pair with the oligonucleotide described in (1) above is not particularly limited, as long as it is able to amplify the sequence of interest in a gene used as a target in PCR, together with the oligonucleotide used in detection of gene polymorphism described in (1) above, in the nucleotide sequence of a gene used as a target for detection of gene polymorphism. Specifically, such a nucleotide sequence can be any given partial sequence consisting of 15 to 40, preferably 18 to 35, and more preferably 18 to 25 continuous nucleotides in a sequence which, rather than being at a position close

to the 5'-end of the sequence which is complementary to the X primer, exists in a region towards the 5'-end of the strand. However, if an oligonucleotide used in detection of gene polymorphism and an oligonucleotide used therewith have sequences complementary to each other, they will anneal with each other, so that a non-specific sequence is amplified, thereby causing a risk of preventing detection of a specific gene polymorphism. Thus, it is preferable that the oligonucleotides be designed so as to avoid such a combination. [0044]

In the present specification, such an oligonucleotide used as part of a pair may be called a reverse primer. [0045]

(b) TaqMan probe

The 5'-end of an oligonucleotide used in detection of gene polymorphism in TaqMan PCR (TaqMan probe) is labeled with a fluorescent reporter dye such as FAM or VIC, and the 3'-end thereof is labeled with a quencher [Genet. Anal., 14, pp.143-149 (1999); J. Clin. Microbiol.], 34, pp. 2933-2936 (1996)].

[0046]

The nucleotide sequences of the oligonucleotide described in (1) above and the TaqMan probe used in a pair with aforementioned nucleotide are not particularly limited, as long as it is able to amplify the sequence of interest in a gene used as the target for detection of gene polymorphism in PCR, together with an oligonucleotide used in detection of gene polymorphism in (1) above. Specifically, such a

nucleotide sequence can be any given partial sequence consisting of 15 to 40, preferably 18 to 35, and more preferably 18 to 25 continuous nucleotides in a sequence which, rather than being at a position close to the 5'-end of the sequence of the strand which is complementary to the X primer, exists in a region towards the 5'-end of the strand. However, if an oligonucleotide used in detection of gene polymorphism and the TaqMan probe have sequences complementary to each other, they will anneal to each other, so that a non-specific sequence is amplified, thereby causing a risk of preventing detection of a specific gene polymorphism. Thus, it is preferable that the above oligonucleotide and the TaqMan probe be designed, so as to avoid such a combination.

[0047]

5. Method for detecting gene polymorphism

A. Detection of gene polymorphism by PCR

(1) PCR

The oligonucleotide used in detection of gene polymorphism, which was designed in the above chapter 4, and an oligonucleotide used as part of a pair with the above oligonucleotide, are used to perform a PCR reaction, so as to detect polymorphism at a certain position of a target gene. Herein, PCR can be carried out by any one of the following combinations: (i) the combination of an oligonucleotide consisting of a nucleotide sequence complementary to the reference sequence with the aforementioned nucleotide; (ii) the combination of an oligonucleotide consisting of a nucleotide sequence complementary to the mutant sequence with

the aforementioned nucleotide; and (iii) the combination of (i) with (ii).

[0048]

Reaction conditions for PCR are not particularly limited, as long as a desired nucleic acid sequence can be amplified under such conditions. Thus, PCR can be carried out under conditions which are generally applied by persons skilled in the art. For example, PCR can be carried out as follows.

[0049]

(a) Nucleic acid synthesizing enzyme

As a nucleic acid synthesizing enzyme, an enzyme can appropriately be selected from DNA polymerase, RNA polymerase and reverse transcriptase, depending on the type of nucleic acid used as a template. Examples of DNA polymerase used herein include Taq DNA polymerase derived from *Thermus aquaticus*, Tth DNA polymerase derived from *Thermus thermophilus*, KOD, Pfu, or Pwo DNA polymerase derived from Pyrococcus, and a mixture consisting of the aforementioned heat-resistant polymerases. However, examples are not limited thereto. Since Tth DNA polymerase also has RT activity, this enzyme is characterized in that it can be used alone when RT-PCR is carried out in a one tube-one step method. Reverse transcriptase means an enzyme capable of reverse transcribing RNA into cDNA. Examples of such reverse transcriptase include reverse transcriptases derived from bird Retroviruses such as Rous associated virus (RAV) or Avian myeloblastosis virus (AMV), reverse transcriptases derived from mouse retroviruses such as Moloney murine leukemia virus

(MMLV), and the aforementioned Tth DNA polymerase. However, examples are not limited thereto.

[0050]

(b) PCR reaction

For example, a PCR reaction is carried out as follows:

[0051]

Example of reaction solution composition

Magnesium chloride 2 to 2.5 mM (preferably 2.5 mM);

1 × PCR buffer (10 mM Tris-HCl (pH 8.3 to pH 9.0 at 25°C (preferably pH 8.3)), 50 mM potassium chloride);

dNTPs 0.2 to 0.25 mM (preferably 0.25 mM);

The oligonucleotide used in detection of gene polymorphism and the oligonucleotide used in a pair with the above nucleotide 0.2 to 0.5 μM (preferably 0.2 μM); and

Taq polymerase 1 to 2.5 units (preferably 2.5 units)

Sterilized water is added to the above solution, so as to adjust a total volume to 80 μl, and the total volume of the solution is then added to the total volume of the reaction solution obtained after completion of the reverse transcription reaction. Thereafter, PCR is initiated.

[0052]

Reaction temperature conditions:

The reaction solution is first heated at 94°C for 2 minutes. Thereafter, a temperature cycle consisting of: 90°C to 95°C (preferably 94°C), 30 seconds; 40°C to 65°C (preferably, up to a temperature that is 20°C lower than the dissociation temperature (Tm) calculated based on the properties of the

primers), 30 seconds; and 70°C to 75°C (preferably 72°C) for 1.5 minutes, is repeated for 28 to 50 cycles (preferably 30 cycles). Thereafter, the reaction solution is cooled to 4°C.

[0053]

(2) Detection of gene polymorphism

After completion of PCR, the reaction solution is subjected to electrophoresis, so as to detect whether or not a band of the size of the sequence of interest has been amplified.

[0054]

(a) The case of using an oligonucleotide having a nucleotide sequence complementary to a reference sequence

When amplification of a sequence of interest has been confirmed as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a reference sequence with an oligonucleotide used in a pair with the aforementioned nucleotide, it can be determined that the oligonucleotide in the polymorphic portion is complementary to the oligonucleotide at the 3'-end of a gene polymorphism detection primer, and that there is no polymorphism.

[0055]

On the other hand, when such a sequence of interest cannot be amplified, it can be determined that the oligonucleotide in the polymorphic portion is not complementary to the oligonucleotide at the 3'-end of the gene polymorphism detection primer, and that there is polymorphism.

[0056]

(b) The case of using an oligonucleotide having a nucleotide sequence complementary to a mutant sequence

When amplification of a sequence of interest has been confirmed as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a mutant sequence with an oligonucleotide used in a pair with the aforementioned nucleotide, it can be determined that the oligonucleotide in the polymorphic portion is complementary to the oligonucleotide at the 3'-end of a gene polymorphism detection primer, and that there is polymorphism.

[0057]

On the other hand, when such a sequence of interest cannot be amplified, it can be determined that the oligonucleotide in the polymorphic portion is not complementary to the oligonucleotide at the 3'-end of the gene polymorphism detection primer, and that there is no polymorphism.

[0058]

(c) The case of using both an oligonucleotide having a nucleotide sequence complementary to a reference sequence and an oligonucleotide having a nucleotide sequence complementary to a mutant sequence

When amplification of a sequence of interest has been confirmed as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a reference sequence with an oligonucleotide used in a pair with the aforementioned nucleotide, but when amplification of a sequence of interest has not been confirmed

as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a mutant sequence with an oligonucleotide used in a pair with the aforementioned nucleotide, it can be determined that there is no polymorphism.

[0059]

On the other hand, when amplification of a sequence of interest has not been confirmed as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a reference sequence with an oligonucleotide used in a pair with the aforementioned nucleotide, but when amplification of a sequence of interest has been confirmed as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a mutant sequence with an oligonucleotide used in a pair with the aforementioned nucleotide, it can be determined that there is polymorphism.

[0060]

When an experiment as described above is carried out using an oligonucleotide that does not contain an ENA oligonucleotide, it can be confirmed that a band appears due to non-complementary primer binding even in nucleic acid acting as a template which does not generally generate such a band. Thus, it can be confirmed that the present method enables detection of gene polymorphism at a higher sensitivity than that of the conventional method. In addition, even when LNA is used instead of the ENA unit, some non-complementary primer binding can be confirmed.

[0061]

Moreover, when an oligonucleotide is used in which the ENA unit is disposed at a position other than the third position from the 3'-end thereof, the precision and sensitivity of detection of gene polymorphism decrease.

[0062]

B. Detection of gene polymorphism by TaqMan PCR

Using the oligonucleotide used in detection of gene polymorphism described in the above "A." section and the TaqMan probe described in the above "4." section, TaqMan PCR is carried out employing an ABI PRISM manufactured by ABI in accordance with protocols included therein, so as to detect gene polymorphism.

[0063]

(a) The case of using an oligonucleotide having a nucleotide sequence complementary to a reference sequence

When amplification of a sequence of interest has been confirmed by an increase in fluorescence intensity as a result of TaqMan PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a reference sequence with a TaqMan probe, it can be determined that the oligonucleotide in the polymorphic portion is complementary to the oligonucleotide at the 3'-end of a gene polymorphism detection primer, and that there is no polymorphism.

[0064]

On the other hand, when such a sequence of interest cannot be amplified, it can be determined that the oligonucleotide in the polymorphic portion is not complementary to the

oligonucleotide at the 3'-end of the gene polymorphism detection primer, and that there is polymorphism.

[0065]

(b) The case of using an oligonucleotide having a nucleotide sequence complementary to a mutant sequence

When amplification of a sequence of interest has been confirmed by an increase in fluorescence intensity as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a mutant sequence with a TaqMan probe, it can be determined that the oligonucleotide in the polymorphic portion is complementary to the oligonucleotide at the 3'-end of a gene polymorphism detection primer, and that there is polymorphism.

[0066]

On the other hand, when such a sequence of interest cannot be amplified, it can be determined that the oligonucleotide in the polymorphic portion is not complementary to the oligonucleotide at the 3'-end of the gene detection polymorphism primer, and that there is no polymorphism.

[0067]

(c) The case of using both an oligonucleotide having a nucleotide sequence complementary to a reference sequence and an oligonucleotide having a nucleotide sequence complementary to a mutant sequence

When amplification of a sequence of interest has been confirmed by an increase in fluorescence intensity as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a reference sequence

with a TaqMan probe, but when amplification of a sequence of interest has not been confirmed as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a mutant sequence with a TaqMan probe, it can be determined that there is no polymorphism.

[0068]

On the other hand, when amplification of a sequence of interest has not been confirmed as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a reference sequence with a TaqMan probe, but when amplification of a sequence of interest has been confirmed by an increase in fluorescence intensity as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a mutant sequence with a TaqMan probe, it can be determined that there is no polymorphism.

[0069]

C. Detection of gene polymorphism by a MALDI-TOF/MS method

By partially modifying the method described in "A method for detecting polymorphism by MALDI-TOF/MS method", ("SNP *Idenshi Takei no Senryaku* (Strategy of SNP Gene Polymorphism)," (edited by Yusuke Nakamura), Nakayama Shoten, Tokyo, (2000), pp. 106-117), gene polymorphism can be detected. This method is specifically described below.

[0070]

A PCR product containing a polymorphic portion is amplified from genomic DNA. In this reaction, bases

corresponding to the polymorphic portion and PCR primers are designed such that they do not overlap with each other.

[0071]

Subsequently, dNTP remaining in the PCR reaction system and an oligonucleotide used as a primer are eliminated, so as to obtain a purified PCR product.

[0072]

The purified PCR product is used as a template. The oligonucleotide described in the section: "(1) Oligonucleotides used in detection of gene polymorphism" in the above chapter "4." is added to the above template in excess, such as 10 times or greater, and they are then annealed at a temperature between 90°C and 95°C, followed by a thermal cycle reaction. The type of thermal cycle reaction is not particularly limited, as long as elongation of the oligonucleotide can be confirmed therein. For example, the reaction can be repeated 25 times between two temperatures, 94°C and 37°C, so as to obtain a suitable elongation efficiency.

[0073]

The elongation reaction product obtained is purified, so as to remove salts, buffer, surfactant, and protein. The purified product is spotted on a MALDI plate, and the mass thereof is then analyzed by MALDI-TOF/MS.

[0074]

When the polymorphic portion of a target gene is an oligonucleotide that is complementary to an oligonucleotide used in detection of gene polymorphism, it is confirmed that an elongation reaction product that is formed by adding ddNTP

to the oligonucleotide used in detection of gene polymorphism is amplified. However, when the polymorphic portion is not complementary to the oligonucleotide used in detection of gene polymorphism, such an elongation reaction product is not amplified.

[0075]

If an elongation reaction product is confirmed when an oligonucleotide consisting of a nucleotide sequence complementary to the reference sequence is used, it can be determined that the polymorphic portion is a reference nucleotide and that it does not have gene polymorphism. If such an elongation reaction product is not confirmed, it can be determined that the polymorphic portion is a mutant nucleotide and that it has gene polymorphism.

[0076]

If an elongation reaction product is confirmed when an oligonucleotide consisting of a nucleotide sequence complementary to the mutant sequence is used, it can be determined that the polymorphic portion is a mutant nucleotide and that it has gene polymorphism. If such an elongation reaction product is not confirmed, it can be determined that the polymorphic portion is a reference nucleotide and that it does not have gene polymorphism.

[0077]

In addition, it is also possible to measure a PCR product by detecting the presence or absence of the generated PCR product, using the Qiagen LightCycler system and applying

it to a kit for detecting such a PCR product (e.g. Quantitect SYBR Green PCR kit).

[0078]

6. Confirmation of the existence of gene polymorphism

The method of the present invention enables determination regarding whether polymorphism in nucleic acid used as a template is present in a hetero state or a homo state. Specifically, it can be determined by any one of the methods described in (a) to (c) below.

[0079]

(a) The case of using an oligonucleotide having a nucleotide sequence complementary to a reference sequence

In the case where the amount of a band of interest seen is approximately half of that of an analyte which is known to be in a homo state, when amplification of a sequence of interest has been confirmed by PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a reference sequence with an oligonucleotide used in a pair with the aforementioned nucleotide, it can be determined that the polymorphism is in a hetero state consisting of a reference nucleotide and a mutant nucleotide.

[0080]

(b) The case of using an oligonucleotide having a nucleotide sequence complementary to a mutant sequence

In the case where the amount of a band of interest seen is approximately half of that of an analyte which is known to be in a homo state, when amplification of a sequence of interest has been confirmed by PCR with the combined use of

an oligonucleotide having a nucleotide sequence complementary to a mutant sequence with an oligonucleotide used in a pair with the aforementioned nucleotide, it can be determined that the polymorphism is in a hetero state consisting of a reference nucleotide and a mutant nucleotide.

[0081]

(c) The case of using both an oligonucleotide having a nucleotide sequence complementary to a reference sequence and an oligonucleotide having a nucleotide sequence complementary to a mutant sequence

In the case where amplification of a sequence of interest can be confirmed as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a reference sequence with an oligonucleotide used in a pair with the aforementioned nucleotide, and where amplification of a sequence of interest can also be confirmed as a result of PCR with the combined use of an oligonucleotide having a nucleotide sequence complementary to a mutant sequence with an oligonucleotide used in a pair with the aforementioned nucleotide, it can be determined that polymorphism is in a hetero state.

[0082]

7. Kit for detecting gene polymorphism

Primers and reagents used to carry out the methods of the present invention can be provided as a kit for detecting gene polymorphism. Such a kit may comprise the following items.

Kit 1:

(a) Oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other oligonucleotides are natural oligonucleotides, the 3'-end nucleotide thereof is a nucleotide complementary to the nucleotide of the reference sequence of the polymorphic site of a target gene, and at the other positions the nucleotides are complementary to the nucleotide sequence of the target gene;

(b) Primer capable of amplifying a sequence of interest, together with the oligonucleotide described in (a) above;

(c) DNA polymerase; and

(d) PCR buffer.

[0083]

Kit 2:

(a) Oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other oligonucleotides are natural oligonucleotides, the 3'-end nucleotide thereof is a nucleotide complementary to the nucleotide of the mutant sequence of the polymorphic site of a target gene, and at the other positions the nucleotides are complementary to the nucleotide sequence of the target gene;

(b) Primer capable of amplifying a sequence of interest, together with the oligonucleotide described in (a) above;

(c) DNA polymerase; and

(d) PCR buffer.

[0084]

The kit of the present invention may also comprise various types of reagents used in electrophoresis, dNTPs, a marker used in electrophoresis, etc., as necessary.

[Examples]

[0085]

The present invention will be more specifically described in the following examples, reference examples and test examples. However, these examples are not intended to limit the scope of the present invention. In the following examples, each genetic manipulation technique is carried out using a method described in Molecular Cloning, Sambrook, J., Fritsch, E. F. and Maniatis, T., Cold Spring Harbor Laboratory Press, 1989, or when commercially available reagents or kits are used, they are used in accordance with instructions included therewith, unless otherwise specified.

[0086]

(Example 1)

Synthesis of

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{p}\text{-}5C^{e2p}\text{-}T^{p}\text{-}C^{t}$

Using an automated nucleic acid synthesizer (ABI model 394 DNA/RNA synthesizer, manufactured by Perkin Elmer), the program was carried out at a scale of 40 nmol. With regard to the concentrations of a solvent, reagent and phosphoramidite in each synthesis cycle, the same concentrations as those for synthesis of a natural oligonucleotide were applied. Approximately 0.1 μmol of CPG was used. As a non-natural phosphoramidite, the compound

described in Example 22 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene 4-N-benzoyl-5-methylcytidine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used. A protected oligonucleotide analogue having a sequence of interest was treated with concentrated ammonia water, so as to separate the oligomer from the support and to remove a cyanoethyl group as a protecting group on a phosphorus atom and a protecting group on a nucleobase. The solvent was evaporated under reduced pressure and the residue was purified by reverse phase HPLC (LC-10VP manufactured by Shimadzu Corporation, column: Merck, Chromolith Performance RP-18e (4.6 × 100 mm), solution A: 5% acetonitrile, 0.1 M triethylamine acetate aqueous solution (TEAA), pH 7.0; solution B: acetonitrile, B%: 10% → 50% (10 min, linear gradient); 60°C; 2 ml/min; 254 nm), so as to collect the peak of a product of interest having a dimethoxytrityl group. Thereafter, water was added thereto, and the mixture was then concentrated under reduced pressure, so as to remove TEAA. Thereafter, an 80% acetic acid aqueous solution (200 μl) was added and the mixture was then left for 20 minutes, to deprotect the dimethoxytrityl group. The solvent was distilled away, and the residue was purified by reverse phase HPLC (LC-10VP manufactured by Shimadzu Corporation, column: Merck, Chromolith Performance RP-18e (4.6 × 100 mm); solution A: 5% acetonitrile, 0.1 M TEAA, pH 7.0; solution B: 25% acetonitrile, 0.1 M TEAA, B%: 0% → 40% (10 min, linear gradient); 60°C; 2 ml/min; 254 nm), to collect the peak of a product of interest. The solvent was distilled

away under reduced pressure and the residue was dissolved in 1 ml of water (9.4 A_{260} units). In addition, the present compound was identified by negative-ion ESI mass spectrometry (calculated value: 6214.11, measurement value: 6214.62).

[0087]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of a human prothrombin gene (GenBank accession No. M17262).

[0088]

(Example 2)

Synthesis of

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{p}\text{-}5C^{e2p}\text{-}T^{p}\text{-}T^{t}$

The compound of Example 2 was synthesized in the same manner as described in Example 1 (21 A_{260} units). The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 6229.12, measurement value: 6229.21).

[0089]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of a human prothrombin gene (GenBank accession No. M17262), wherein G is mutated to A at nucleotide No. 26784.

[0090]

(Example 3)

Synthesis of

HO-A^p-T^p-C^p-T^p-G^p-T^p-C^p-T^p-A^p-C^p-A^p-T^p-A^p-T^p-A^p-T^p-A^p-T^p-A^p-C^p-A^p-C^p-A^p-C^p-A^p-$5C^{e2p}$-A^p-T^t

The compound of Example 3 was synthesized in the same manner as described in Example 1 (8.9 A_{260} units). The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 8530.67, measurement value: 8530.75).

[0091]

The nucleotide sequence of the present compound corresponds to nucleotide Nos. 60529-60556 described in GenBank accession No. AL935325, wherein C is converted to T at nucleotide No. 60556.

[0092]

(Example 4)

Synthesis of

HO-A^p-T^p-C^p-T^p-G^p-T^p-C^p-T^p-A^p-C^p-A^p-T^p-A^p-T^p-A^p-T^p-A^p-T^p-A^p-C^p-A^p-C^p-A^p-C^p-A^p-$5C^{e2p}$-A^p-C^t

The compound of Example 4 was synthesized in the same manner as described in Example 1 (10.1 A_{260} units). The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 8515.66, measurement value: 8515.56).

[0093]

The nucleotide sequence of the present compound corresponds to nucleotide Nos. 60529-60556 described in GenBank accession No. AL935325.

[0094]

(Reference Example 1)

HO-C^p-A^p-C^p-T^p-G^p-G^p-G^p-A^p-G^p-C^p-A^p-T^p-T^p-G^p-A^p-G^p-G^p-C^p-T^p-C^t

The compound of Reference Example 1 was purchased from Sigma-Genosys Japan KK.

[0095]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of a human prothrombin gene (GenBank accession No. M17262), and it is shown in SEQ ID NO: 1 of the sequence listing.

[0096]

(Reference Example 2)

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{p}\text{-}C^{p}\text{-}T^{p}\text{-}T^{t}$

The compound of Reference Example 2 was purchased from Sigma-Genosys Japan KK.

[0097]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of a human prothrombin gene (GenBank accession No. M17262), wherein G is mutated to A at nucleotide No. 26784. This sequence is shown in SEQ ID NO: 2 of the sequence listing.

[0098]

(Reference Example 3)

Synthesis of

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{p}\text{-}C^{p}\text{-}T^{p}\text{-}C^{e2t}$

The compound of Reference Example 3 was synthesized in the same manner as described in Example 1 (0.3 A_{260} units). However, as non-natural phosphoramidite, the compound described in Example 5 of Japanese Patent No. 3420984

(5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-4-N-benzoylcytidine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used and as a solid support, approximately 0.1 μmol of universal-Q 500 CPG (manufactured by Glen Research) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 6200.08, measurement value: 6200.25).

[0099]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of the human prothrombin gene (GenBank accession No. M17262).

[0100]

(Reference Example 4)

Synthesis of

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{p}\text{-}C^{p}\text{-}T^{p}\text{-}T^{e2t}$

The compound of Reference Example 4 was synthesized in the same manner as described in Example 1 (0.94 A_{260} units). However, as non-natural phosphoramidite, the compound described in Example 9 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-5-methyluridine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used and as a solid support, approximately 0.1 μmol of universal-Q 500 CPG (manufactured by Glen Research) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 6215.09, measurement value: 6215.06).

[0101]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of the human prothrombin gene (GenBank accession No. M17262), wherein G is mutated to A at nucleotide No. 26784.

[0102]

(Reference Example 5)

Synthesis of

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{p}\text{-}C^{p}\text{-}T^{e2p}\text{-}C^{t}$

The compound of Reference Example 5 was synthesized in the same manner as described in Example 1 (2.28 A_{260} units). However, as non-natural phosphoramidite, the compound described in Example 9 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-5-methyluridine -3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 6200.08, measurement value: 6200.26).

[0103]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of the human prothrombin gene (GenBank accession No. M17262).

[0104]

(Reference Example 6)

Synthesis of

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{p}\text{-}C^{p}\text{-}T^{e2p}\text{-}T^{t}$

The compound of Reference Example 6 was synthesized in the same manner as described in Example 1 (2.28 A_{260} units). However, as non-natural phosphoramidite, the compound described in Example 9 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-5-methyluridine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 6215.09, measurement value: 6215.26).

[0105]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of the human prothrombin gene (GenBank accession No. M17262), wherein G is mutated to A at nucleotide No. 26784.

[0106]

(Reference Example 7)

Synthesis of

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{e2p}\text{-}C^{p}\text{-}T^{p}\text{-}C^{t}$

The compound of Reference Example 7 was synthesized in the same manner as described in Example 1 (4.32 A_{260} units). However, as non-natural phosphoramidite, the compound described in Example 27 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-2-N-isobutyrylguanosine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 6200.08, measurement value: 6199.95).

[0107]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of the human prothrombin gene (GenBank accession No. M17262).

[0108]

(Reference Example 8)

Synthesis of

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{e2p}\text{-}C^{p}\text{-}T^{p}\text{-}T^{t}$

The compound of Reference Example 8 was synthesized in the same manner as described in Example 1 (8.0 A_{260} units). However, as non-natural phosphoramidite, the compound described in Example 27 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-2-N-isobutyrylguanosine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 6215.09, measurement value: 6215.06).

[0109]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of the human prothrombin gene (GenBank accession No. M17262), wherein G is mutated to A at nucleotide No. 26784.

[0110]

(Reference Example 9)

Synthesis of

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{p}\text{-}C^{e1p}\text{-}T^{p}\text{-}C^{t}$

The compound of Reference Example 9 was synthesized in the same manner as described in Example 1 (13.28 A_{260} units). However, as non-natural phosphoramidite, the compound (C^{elp}), 5'-O-dimethoxytrityl-2'-O,4'-C-methylene-4-N-benzoylcytidine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite, described in the publication, Tetrahedron (1998) 54, 3607-3630, was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 6186.05, measurement value: 6186.45).

[0111]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of the human prothrombin gene (GenBank accession No. M17262).

[0112]

(Reference Example 10)

Synthesis of

$HO\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}G^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}T^{p}\text{-}G^{p}\text{-}A^{p}\text{-}G^{p}\text{-}G^{p}\text{-}C^{elp}\text{-}T^{p}\text{-}T^{t}$

The compound of Reference Example 10 was synthesized in the same manner as described in Example 1 (8.0 A_{260} units). However, as non-natural phosphoramidite, the compound (C^{elp}), 5'-O-dimethoxytrityl-2'-O,4'-C-methylene-4-N-benzoylcytidine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite, described in the publication, Tetrahedron (1998) 54, 3607-3630, was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 6201.07, measurement value: 6201.14).

[0113]

The nucleotide sequence of the present compound is complementary to a nucleotide sequence corresponding to nucleotide Nos. 26784-26803 of the human prothrombin gene (GenBank accession No. M17262), wherein G is mutated to A at nucleotide No. 26784.

[0114]

(Reference Example 11)

Synthesis of

$HO\text{-}A^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{t}$

The compound of Reference Example 11 was purchased from Sigma-Genosys Japan KK.

[0115]

The nucleotide sequence of the present compound corresponds to nucleotide Nos. 60529-60556 described in GenBank accession No. AL935325, wherein C is converted to T at nucleotide No. 60556. This sequence is shown in SEQ ID NO: 3 of the sequence listing.

[0116]

(Reference Example 12)

Synthesis of

$HO\text{-}A^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{t}$

The compound of Reference Example 12 was purchased from Sigma-Genosys Japan KK.

[0117]

The nucleotide sequence of the present compound corresponds to nucleotide Nos. 60529-60556 described in

GenBank accession No. AL935325. This sequence is shown in SEQ ID NO: 4 of the sequence listing.

[0118]

(Reference Example 13)

Synthesis of

HO-A^p-T^p-C^p-T^p-G^p-T^p-C^p-T^p-A^p-C^p-A^p-T^p-A^p-T^p-A^p-T^p-A^p-T^p-A^p-C^p-A^p-C^p-A^p-C^p-A^p-C^p-A^p-T^{e2t}

The compound of Reference Example 13 was synthesized in the same manner as described in Example 1 (7.8 A_{260} units). However, as non-natural phosphoramidite, the compound described in Example 9 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-5-methyluridine -3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used, and as a solid support, approximately 0.1 μmol of universal-Q 500 CPG (manufactured by Glen Research) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 8516.64, measurement value: 8515.88).

[0119]

The nucleotide sequence of the present compound corresponds to nucleotide Nos. 60529-60556 described in GenBank accession No. AL935325, wherein C is converted to T at nucleotide No. 60556.

[0120]

(Reference Example 14)

Synthesis of

HO-A^p-T^p-C^p-T^p-G^p-T^p-C^p-T^p-A^p-C^p-A^p-T^p-A^p-T^p-A^p-T^p-A^p-T^p-A^p-C^p-A^p-C^p-A^p-C^p-A^p-C^p-A^p-$5C^{e2t}$

The compound of Reference Example 14 was synthesized in the same manner as described in Example 1 (7.4 A_{260} units). However, as a solid phase carrier, approximately 0.1 μmol of universal-Q 500 CPG (manufactured by Glen Research) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 8516.66, measurement value: 8516.00).

[0121]

The nucleotide sequence of the present compound corresponds to nucleotide Nos. 60529-60556 described in GenBank accession No. AL935325.

[0122]

(Reference Example 15)

Synthesis of

$HO\text{-}A^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{e2p}\text{-}T^{t}$

The compound of Reference Example 15 was synthesized in the same manner as described in Example 1 (8.4 A_{260} units). However, as non-natural phosphoramidite, the compound described in Example 14 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-6-N-benzoyladenosine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 8516.64, measurement value: 8516.32).

[0123]

The nucleotide sequence of the present compound corresponds to nucleotide Nos. 60529-60556 described in

GenBank accession No. AL935325, wherein C is converted to T at nucleotide No. 60556.

[0124]

(Reference Example 16)

Synthesis of

$HO\text{-}A^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{e2p}\text{-}C^{t}$

The compound of Reference Example 16 was synthesized in the same manner as described in Example 1 (7.9 A_{260} units). However, as non-natural phosphoramidite, the compound described in Example 14 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-6-N-benzoyladenosine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 8501.63, measurement value: 8500.70).

[0125]

The nucleotide sequence of the present compound corresponds to nucleotide Nos. 60529-60556 described in GenBank accession No. AL935325.

[0126]

(Reference Example 17)

Synthesis of

$HO\text{-}A^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{e2p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{t}$

The compound of Reference Example 17 was synthesized in the same manner as described in Example 1 (9.7 A_{260} units). However, as non-natural phosphoramidite, the compound

described in Example 14 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-6-N-benzoyladenosine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 8516.64, measurement value: 8517.14).

[0127]

The nucleotide sequence of the present compound corresponds to nucleotide Nos. 60529-60556 described in GenBank accession No. AL935325, wherein C is converted to T at nucleotide No. 60556.

[0128]

(Reference Example 18)

Synthesis of

$HO\text{-}A^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}G^{p}\text{-}T^{p}\text{-}C^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}T^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{p}\text{-}A^{e2p}\text{-}C^{p}\text{-}A^{p}\text{-}C^{t}$

The compound of Reference Example 18 was synthesized in the same manner as described in Example 1 (7.2 A_{260} units). However, as non-natural phosphoramidite, the compound described in Example 14 of Japanese Patent No. 3420984 (5'-O-dimethoxytrityl-2'-O,4'-C-ethylene-6-N-benzoyladenosine-3'-O-(2-cyanoethyl N,N-diisopropyl)phosphoramidite) was used. The present compound was identified by negative-ion ESI mass spectrometry (calculated value: 8501.63, measurement value: 8501.65).

[0129]

The nucleotide sequence of the present compound corresponds to nucleotide Nos. 60529-60556 described in GenBank accession No. AL935325.

[0130]

(Test example 1) Detection of SNP in human prothrombin gene

In order to detect SNP (F2 20210G-A) in a human prothrombin gene (coagulation factor II; GenBank accession No. M17262), a reverse primer and human DNA were prepared using TrueSNP Demo Kit (Proligo) in accordance with the protocols included therein. The nucleotide sequence of the reverse primer corresponds to nucleotide Nos. 26588-26605 of GenBank accession No.M17262. The sequence is as follows: 5'-GGGTGAAGGCTGTGACCG-3' (SEQ ID NO: 5 of the sequence listing)

As a forward primer, the compound (1.25 μM) described in any one of Examples 1 and 2, and Reference Examples 1 to 8, was used. A solution comprising 5 μl of the forward primer, 1.3 μl of the reverse primer, 12.5 μl of Premix Taq (manufactured by Takara Shuzo Co., Ltd.), 1 μl of a human DNA solution, and 5.2 μl of sterilized water, was subjected to a PCR reaction, using a Takara PCR Thermal Cycler PERSONAL (TP240). As the reaction cycle, after a treatment at 94°C for 10 minutes, a cycle consisting of 94°C, 1 minute, 63°C, 1 minute, and 72°C, 1 minute, was repeated for 31 cycles. After completion of the reaction, 1 μl of a loading solution was added to 5 μl of the reaction solution, this was followed by 10%

polyacrylamide gel electrophoresis (1 × TBE, 200 V constant voltage, approximately 1 hour). Thereafter, the resultant gel was stained with SYBR Green I (manufactured by Cambrex), and the Molecular Imager FX Fluorescent Imager system (Bio-Rad) was used to visualize the band. Thereafter, it was quantified using Quantity One software (Bio-Rad).

[0131]

The results are shown in Figure 3. When a compound, wherein an ENA unit had been introduced into the third position from the 3'-end thereof, was used as a forward primer, in the case of the compound of Example 1, amplification of the gene of interest (216 bp) could be confirmed, but in the case of the compound of Example 2, amplification of the gene of interest (216 bp) could not be confirmed. In contrast, when the compounds of Reference Examples 1 and 2, which were natural oligonucleotides, were used as forward primers, not only in the case of the compound of Reference Example 1, but also in the case of the compound of Reference Example 2, amplification of the gene of interest was confirmed, and thus amplification of the gene due to non-complementary binding took place. On the other hand, in the case of the compounds of Reference Examples 3 and 4, wherein an ENA unit had been introduced into the 3'-end thereof, and in the case of the compounds of Reference Examples 5 and 6, wherein such an ENA unit had been introduced into the second position from the 3'-end thereof, amplification of a gene of interest was not confirmed. These results revealed that when a compound, wherein an ENA unit has been introduced into the third position

from the 3'-end thereof, is used as a primer, there is almost no non-complementary binding, thereby selectively amplifying the gene (216 bp).

[0132]

Figure 4 shows an example using Premix EX Taq (manufactured by Takara Shuzo Co., Ltd.) instead of Premix Taq (manufactured by Takara Shuzo Co., Ltd.). In this case also, when a primer wherein an ENA unit had been introduced into the third position from the 3'-end thereof was used, almost no non-complementary binding took place, and when the primer of Example 1 was used, the gene was amplified more efficiently and selectively.

[0133]

The fluorescence intensity of the detected band was converted into a numerical value, and it was then plotted, as shown in Figure 5. In the case of the compounds of Reference Examples 9 and 10, an LNA unit was introduced into the third position from the 3'-end thereof. When the compound of Reference Example 10 was used as the forward primer, 15% amplification of the gene due to non-complementary binding was observed. In contrast, when the compound of Example 2, wherein an ENA unit had been introduced into the third position from the 3'-end thereof, was used as a forward primer, only 6% amplification of the gene due to non-complementary binding was observed. Thus, it was revealed that such an ENA unit results in little non-complementary binding, having high selectivity.

[0134]

(Test example 2) Detection of polymorphism in angiopoietin-like protein 3 gene promoter

(1) Preparation of mouse genomic DNA

A tail tip (1.5 cm) collected from each of the mice (mouse AKR strain, KK mouse Nga strain, and KK mouse Snk strain (4-week-old)) was immersed in 840 µl of a dissolving solution (consisting of 720 µl of 1 × SSC, 80 µl of 10% SDS, and 40 µl of 10 mg/ml proteinase K) and the samples were then shaken overnight while being incubated at 50°C. Subsequently, 20 µl of 1 mg/ml ribonuclease A was added to the reaction solution, and it was then incubated at 50°C for 1 hour. Thereafter, phenol-chloroform extraction was carried out twice, and an ethanol precipitation operation was then carried out once. The precipitate was dissolved in 150 µl of a buffer comprising 10 mM Tris-HCl (pH 7.5) and 1 mM EDTA. Thereafter, the solution was subjected to spectrophotometry (U-3000, manufactured by Hitachi, Ltd.), so as to measure the absorbance at a wavelength of 260 nm. Thereafter, sterilized water was added thereto so as to adjust the concentration to 25 ng/µl, thereby preparing a genomic DNA sample.

[0135]

(2) PCR

From the results of direct sequencing, polymorphism in the angiopoietin-like protein 3 gene promoter is as shown in Figure 6. Figure 6 shows that when compared with the mouse KK/Nga strain and the KK/Snk strain, the mouse AKR strain has polymorphism, such that 2 bases (CA) indicated with the mark " : " are deleted.

As the reverse primer for PCR, a primer having the following sequence was used:

5'-GTCACTAGACTACTGCTTACTGTCC-3' (SEQ ID NO: 6 of the sequence listing)

(The nucleotide sequence of the present compound is complementary to a sequence corresponding to nucleotide Nos. 60658-60682 described in GenBank accession No.AL935325). As the forward primer, a compound (1.25 μM) described in any of Examples 3 and 4 and Reference Examples 11 to 18 was used. A solution comprising 5 μl (1.25 μM) of the forward primer, 5 μl (1.25 μM) of the reverse primer, 12.5 μl of Premix Taq (manufactured by Takara Shuzo Co., Ltd.), 0.125 μl of the genomic DNA solution (100 ng/1 μl) and 2.38 μl of sterilized water was subjected to a PCR reaction, using a Takara PCR Thermal Cycler PERSONAL (TP240). As the reaction cycle, after a heat treatment at 94°C for 10 minutes, a cycle consisting of 94°C, minute, 63°C, 1 minute, and 72°C, 1 minute, was repeated for 30 cycles. After completion of the reaction, 1 μl of a loading solution was added to 5 μl of the reaction solution, this was followed by 10% polyacrylamide gel electrophoresis (1 × TBE, 200 V constant voltage, approximately 1 hour). Thereafter, the resultant gel was stained with SYBR Green I (manufactured by Cambrex), and the Molecular Imager FX Fluorescent Imager system (Bio-Rad) was used to visualize the band.

[0136]

(3) Results

The results obtained using the mouse AKR strain-derived genomic DNA (AKR) as the template are shown in Figure 7-A. It was found that in the compounds of Examples 3 and 4, wherein an ENA unit has been introduced into the third position from the 3'-end thereof, in the case of using the compound of Example 3 as a forward primer, a gene (152 bp) was selectively amplified.

[0137]

The results obtained using the KK mouse Nga strain-derived genomic DNA (KK/Nga) as the template are shown in Figure 7-B. It was found that in the compounds of Examples 3 and 4, wherein an ENA unit has been introduced into the third position from the 3'-end thereof, in the case of using the compound of Example 4 as the forward primer, the gene was amplified most efficiently and selectively.

[0138]

Figure 8 shows the results of PCR, using the compounds of Example 3 and 4 as forward primers and also using, as a template, AKR genomic DNA, KK/Nga genomic DNA, KK mouse Snk strain (KK/Snk) genomic DNA and DNA (Mix) formed by mixing equal amounts of AKR and KK/Nga genomic DNAs. As shown in Figure 8A, in the case of AKR, when the compound of Example 3 was used as the forward primer, selective amplification of the gene was confirmed. In the case of KK/Nga and KK/Snk, when the compound of Example 4 was used as the forward primer, selective amplification of the gene was confirmed. In the case of Mix, in both cases of using the compound of Example 3 or the compound of Example 4 as the forward primer,

amplification of the gene was confirmed. Thus, it was indicated that even if polymorphism is in a hetero state, it is distinguishable. Moreover, as shown in Figure 8B, when the compounds of Reference Examples 11 and 12, which are natural DNAs, were used as forward primers, by-products were observed in addition to amplification of the band of interest. Thus, it was found that the combination of the compound of Example 3 with that of Example 4, wherein an ENA unit has been introduced into the third position from the 3'-end thereof, is far better in terms of detection of gene polymorphism.

[Brief Description of the Drawings]

[0139]

[Figure 1] Figure 1 is a view showing the principle of a method for detecting gene polymorphism when there is no polymorphism.

[Figure 2] Figure 2 is a view showing the principle of a method for detecting gene polymorphism when there is polymorphism.

[Figure 3] Figure 3 is a view showing the results of PCR using Premix Taq and various types of primers.

[Figure 4] Figure 4 is a view showing the results of PCR using Premix EX Taq and various types of primers.

[Figure 5] Figure 5 is a view showing the fluorescence intensity of each band detected by PCR, which has been converted into numerical values.

[Figure 6] Figure 6 is a view showing polymorphism in an angiopoietin-like protein 3 gene promoter.

[Figure 7-A] Figure 7-A is a view showing the results of PCR using mouse AKR strain-derived genomic DNA (AKR) as a template.

[Figure 7-B] Figure 7-B is a view showing the results of PCR using KK mouse Nga strain-derived genomic DNA (KK/Nga) as a template.

[Figure 8] Figure 8 is a view showing the results of PCR, using, as templates, AKR genomic DNA, KK/Nga genomic DNA, KK mouse Snk strain (KK/Snk) genomic DNA, and DNA formed by mixing equal amounts of AKR and KK/Nga genomic DNAs.

[Title of the Document] Abstract

[Abstract]

[Problem to be Solved]

It is an object of the present invention to provide a method for detecting gene polymorphism and an oligonucleotide used for the above method, and to further provide a kit for detecting gene polymorphism, comprising the above oligonucleotide.

[Solution]

The present invention provides: an oligonucleotide, wherein the third nucleotide from the 3'-end thereof is a 2'-*O*,4'-*C*-ethylene nucleotide, the other oligonucleotides are natural oligonucleotides, the 3'-end position thereof is a nucleotide complementary to the nucleotide of the reference sequence of a polymorphic sequence of a target gene, and the other positions are nucleotides complementary to the nucleotide sequence of the target gene; a method for detecting gene polymorphism using the above oligonucleotide; a kit for detecting gene polymorphism comprising the above oligonucleotide; and the like.

[Selected Drawing] None

SEQUENCE LISTING

<110> SANKYO COMPANY, LIMITED

<120> Method for identifying SNPs

<130> 2003175SU

<160> 6

<170> PatentIn version 3.1

<210> 1
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<223> Inventor: Koizumi, Makoto

<400> 1
cactgggagc attgaggctc 20

<210> 2
<211> 20
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> (20)..(20)
<223>

<220>
<221> allele
<222> (20)..(20)
<223> C is transitonded to T

<400> 2
cactgggagc attgaggctt 20

<210> 3
<211> 28
<212> DNA
<213> Mus musculus

<220>
<221> allele
<222> (28)..(28)
<223> C is transitioned to T

<400> 3
atctgtctac atatatatac acacacat 28

<210> 4
<211> 28
<212> DNA
<213> Mus musculus

<400> 4
atctgtctac atatatatac acacacac 28

<210> 5
<211> 18
<212> DNA
<213> Homo sapiens

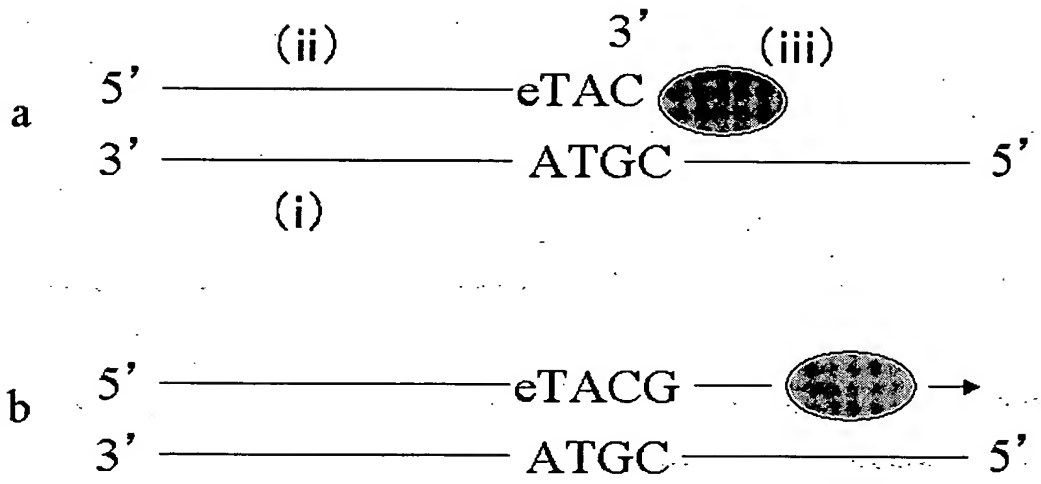
<400> 5
gggtgaaggc tgtgaccg 18

<210> 6
<211> 25
<212> DNA
<213> Mus musculus

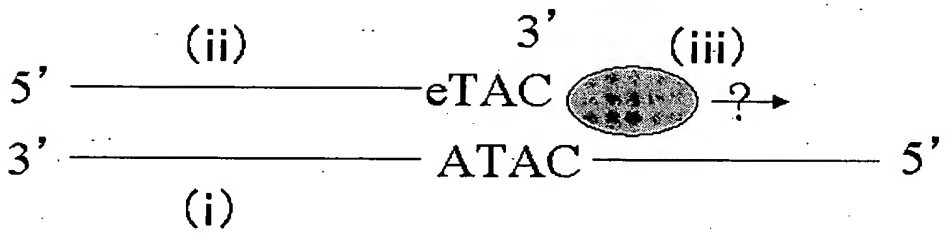
<400> 6
gtcactagac tactgcttac tgtcc 25

[Title of the Document] Figure

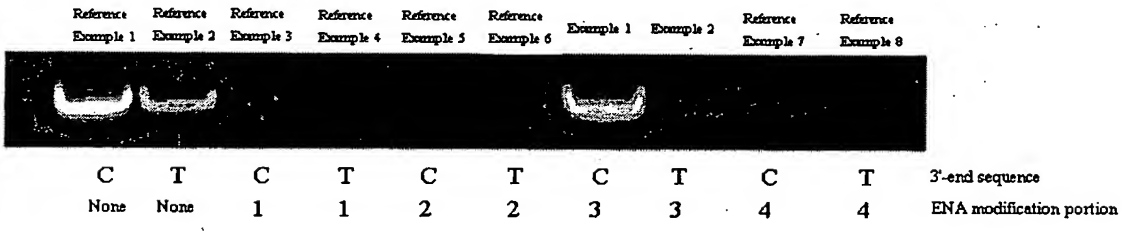
[Figure 1]



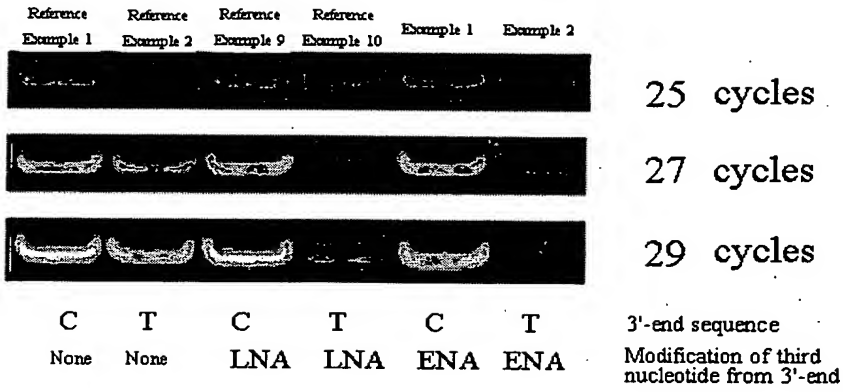
[Figure 2]



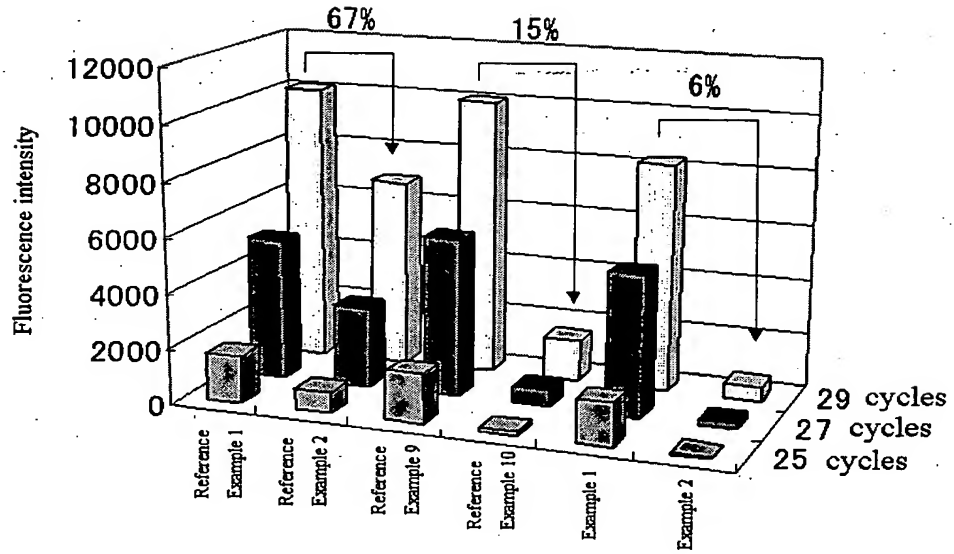
[Figure 3]



[Figure 4]



[Figure 5]

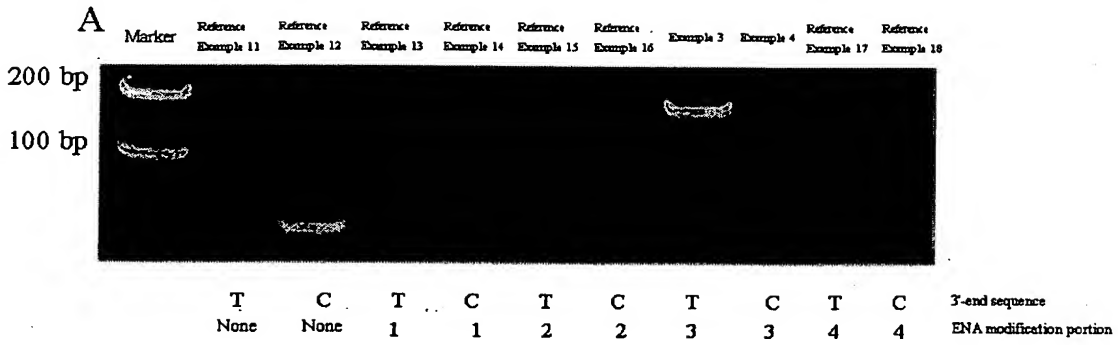


[Figure 6]

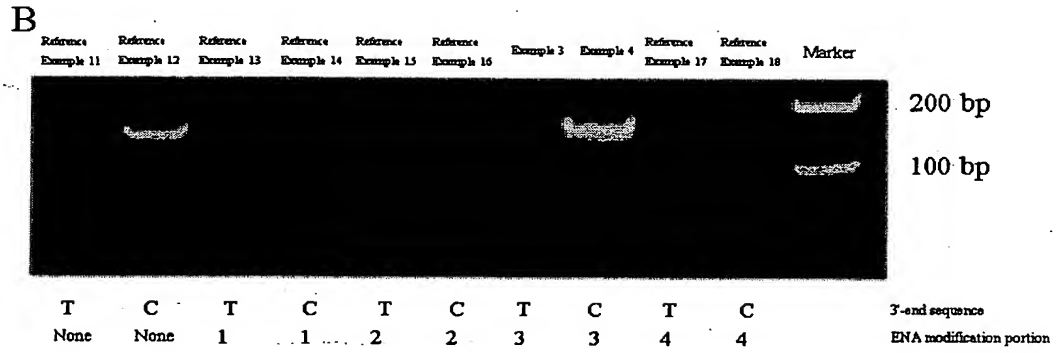
```
(KK/Nga,KK/Snk)                                    ACAT
(AKR)          ---ATCTGTCTACATATATATACACACACA::T---

Primer          5'ATCTGTCTACATATATATACACACACAT3'
                5'ATCTGTCTACATATATATACACACACAC3'
```

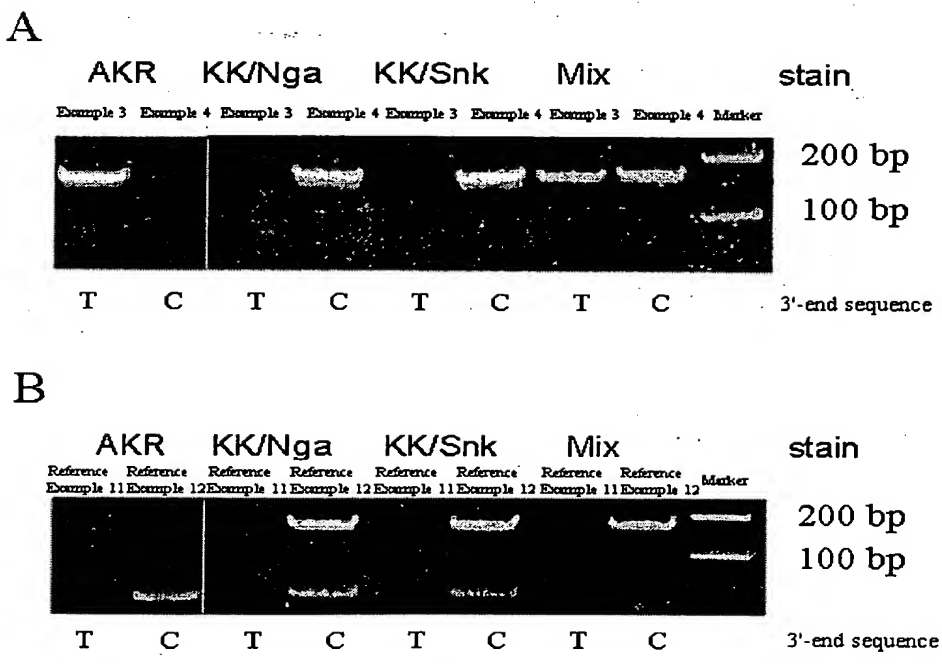
[Figure 7-A]



[Figure 7-B]



[Figure 8]



Acknowledged · Added information

Patent Application No.	2003-378039
Receipt No.	50301845086
Name of Document	Patent Application
Examiner in charge	Haruho Ueda 6992
Date of compiling	November 10, 2003

[Acknowledged information · Added information]

Filing date	November 7, 2003

Information on Applicant's History

Identification Number	[000001856]
1. Date of Alteration	August 15, 1990
[Reason for Alteration]	Newly registered
Address	5-1, Nihonbashi Honcho 3-chome, Chuo-ku, Tokyo, Japan
Name	SANKYO COMPANY, LIMITED

CHROMOLITH® HPLC COLUMNS

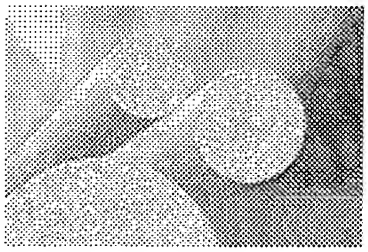
High speed, high performance in monolithic form

Chromatographers today seek ways to analyse more samples – by speeding up the separation process, and therefore analysis. Merck's Chromolith® HPLC Columns provide excellent separations in a fraction of the time taken by standard 5 µm particulate columns, at half the column back pressure.

Typical particulate HPLC columns are filled with silica particles of 3 or 5 microns. Despite its high separation performance, such small particle size limits the speed of analysis and column length, plus reduces column lifetime from blockages and the lifetime of HPLC instrument and injector seals at high pressure.

Benefits at a glance

- Up to 9 times the speed of separations
- Half the column back pressure compared to 5 µm columns
- Significantly increased column lifetime
- Column length no longer pressure limited
- Standard HPLC instruments well suited
- Cost savings from increased sample throughput and column lifetime



Chromolith® HPLC Columns from Merck, however, are not filled with minute silica particles like conventional packed HPLC columns, but consist of a single rod of high purity monolithic silica with a revolutionary bimodal pore structure. Its macropores dramatically reduce the column back pressure to enable faster flow rates, thereby considerably reducing the analysis time. While mesopores form the fine porous structure, and provide a large active surface area for highly efficient separations.

A winning innovation

Winner of numerous awards including one of the five most important innovations of 2001, Merck's Chromolith® HPLC Columns are constantly finding new applications in every area of pharmaceutical and biological research, plus food, chemical and environmental analysis.

Recent examples include analysis of illicit drugs in just half the time of conventional methods, and analysis of pollutant phenols in just three minutes – which previously took 18.
With longer lifetime, less matrix sensitivity with biological samples, and the possibility of coupling multiple columns together for efficiencies of 100.000 plates at normal pressure, Merck Chromolith® HPLC Columns offer speed and performance previously unseen in monolithic form.
If you would like further information, see also our brochure on Chromolith® HPLC Columns or please contact us.

J Hum Genet (2007) 52:871–880
DOI 10.1007/s10038-007-0200-z

MINIREVIEW

SNPs in disease gene mapping, medicinal drug development and evolution

Barkur S. Shastry

Received: 28 July 2007 / Accepted: 18 September 2007 / Published online: 11 October 2007

Abstract Single nucleotide polymorphism (SNP) technologies can be used to identify disease-causing genes in humans and to understand the inter-individual variation in drug response. These areas of research have major medical benefits. By establishing an association between the genetic make-up of an individual and drug response it may be possible to develop a genome-based diet and medicines that are more effective and safer for each individual. Additionally, SNPs can be used to understand the molecular mechanisms of sequence evolution. It has been found that throughout the given gene, the rate, type and site of nucleotide substitutions as well as the selection pressure on codons is not uniform. The residues that evolve under strong selective pressures are found to be significantly associated with human disease. Deleterious mutations that affect biological function of proteins are effectively being rejected by natural selection from the gene pool. If substituted nucleotides are fixed during evolution then they may have selection advantages, they may be neutral, or they may be deleterious and cause pathology. Therefore, it is possible that disease-associated SNPs (or pathology) and evolution can be related to one another.

Keywords Evolution · Genomics · Medicine · Pharmacology · Polymorphism

B. S. Shastry (✉)
Department of Biological Sciences,
Oakland University, Rochester, MI, USA
e-mail: shastry@oakland.edu

Introduction

The completion of the human genome project has generated new enthusiasm and opportunities in life sciences. It has provided the necessary tools to understand the genetic basis of diversity among individuals, the most common familial traits, evolutionary processes, complex and common diseases such as diabetes, obesity, hypertension and psychiatric disorders, and to develop genome-based medicinal drugs (Emilien et al. 2000). Scientists generally think that the genomes between two randomly selected individuals contain approximately 0.1% differences or variations. This variation is called polymorphism, and it arises because of mutations. Several comparative studies on identical and fraternal twins (Martin et al. 1997) and siblings suggest that DNA polymorphism is one of the factors associated with susceptibility to many common diseases (Table 1), every human trait such as curly hair, individuality and inter-individual difference in drug response. DNA sequence variation is also considered to be responsible for genome evolution. Based on these observations, it has been proposed that by cataloging the DNA polymorphisms in different populations and in different species, it may be possible (a) to develop genome-based knowledge on the susceptibility of an individual to many common diseases, (b) to manufacture safer and more effective individualized diet and medications for patients, and (c) to understand evolutionary processes. However, many experts believe that this single nucleotide polymorphism (SNP) technology has to face several challenges before it makes its impact on medicine. What follows is a brief discussion of the above three aspects with emphasis placed on evolution.

Table 1 A partial list of diseases associated with single nucleotide polymorphisms

Disease	Gene	Disease	Gene
Asthma	EDN1 and NOS1	Lung cancer	MMP1
	Chemokine		p53
Arrhythmia	KCN1	Myocardial	TSP
		Infraction	PCS
Blood pressure	TAF1	Migraine	IR
Biliary cirrhosis	MBL	Obesity	PAI1
Bipolar affective disorder	HRT 3A	Ossification	Npps
Colorectal cancer	Cyclin D1	Oxalate stone	E-Cad
Crohn's Disease	MDR1	POAG	Myocilin
Dyslipidemia	Lipase	Rheumatoid arthritis	MIF
Eating disorder	Melanocortin	Systemic sclerosis	Fibrillin1
Esophagel adenocarcinoma	Cyclin D1	Severe sepsis	TNF-α
Hyperbilirubinemia	UGT1A1	Type II diabetes	Syntaxin1A
Idiopathic arthritis	MIF	Ulcerative colitis	MDR1
Idiopathic PD and FTD	Tau	Urinary bladder cancer	Cyclin D1
Knee and hip osteoarthritis	Collagen	Autism	CNP

EDN1 endothelin 1, *NOS1* neuronal nitric oxide synthetase 1, *KCN1* potassium channel protein, *TAF1* thrombin-activatable fibrinolysis inhibitor, *MBL* mannose binding protein, *UGT1A1* UDP glucoronosyl transferase, *MIF* macrophage migration inhibitory factor, *PD* Parkinson disease, *FTD* frontotemporal dementia, *MMP1* matrix metalloproteinase 1, *PAI* plasminogen activator inhibitor, *Npps* nucleotide pyrophosphatase, *Cad* cadherin, *POAG* primary open angle glaucoma, *TNF* tumor necrosis factor, *MDR1* p-glycoprotein (multiple-drug-resistant), *TSP* thrombospondin, *PCS* prostacyclin synthase, *IR* insulin receptor, *CNP* copy number polymorphism

Detection and analysis of DNA polymorphism

The simplest form of DNA variation among individuals is the substitution of one single nucleotide for another. This type of change (Fig. 1A) is called SNP. It is estimated that SNPs occur at a frequency of 1 in 1,000 bp throughout the genome. These simple changes can be of transition or transversion type. According to one report (Halushka et al. 1999), approximately 50% of SNPs are in the noncoding regions, 25% lead to missense mutations (coding SNPs or cSNPs), and the remaining 25% are silent mutations (they do not change encoded amino acids). These silent SNPs are called synonymous SNPs, and it is most likely that they are not subject to natural selection (but see below). On the other hand, nonsynonymous SNPs (nSNPs, change-encoded amino acids) may produce pathology and may be subject to natural selection. SNPs (both synonymous and nonsynonymous) influence promoter activity and pre-mRNA conformation (or stability). They also alter the ability of a protein to bind its substrate or inhibitors (Kimchi-Sarfaty et al. 2007) and change the subcellular localization of proteins (nSNPs). Therefore, they may be responsible for disease susceptibility, medicinal drug deposition and genome evolution. Although several of them affect the functions of genes, many of them are not deleterious to organisms and must have escaped selection pressure. For the purpose of identifying SNPs, several private and public organizations

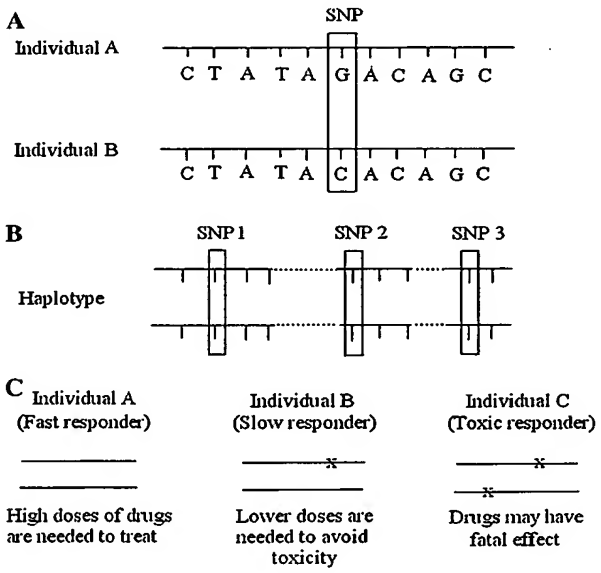


Fig. 1A–C A schematic representation of single nucleotide polymorphism (*SNP*) (**A**), a haplotype (**B**), and the relationship between the genotype and variation in drug response among three individuals (**C**). In **A**, strings of nucleotides at which individuals *A* and *B* differ are shown. In **B**, a long stretch of DNA with distinctive patterns of SNPs at a given location of a chromosome is shown. Haplotype diversity may be generated by new SNP alleles. In **C**, two *horizontal lines* denote a pair of homologous genes and the symbol *X* indicates polymorphism in the gene

have undertaken massive efforts to develop high-throughput SNP genotyping methods over the past 20 years (reviewed in Shastry 2002, 2005). As a result of these efforts, a large collection of SNPs is now available from the human genome project (http://www.ncbi.nlm.nih.gov/SNP/snp_summary.cgi).

SNPs in gene discovery

Because some diseases are hereditary, one immediate goal of the human genome project is to find out which genes predispose people to various disorders and how the sequence variation in a gene affects the functions of its product. As mentioned previously, SNPs occur frequently throughout the genome. Therefore, they can be used as markers to identify disease-causing genes by an association study (Gray et al. 2000). In such studies, it is assumed that two closely located alleles (gene and marker) are inherited together. Therefore, a simple comparison of patterns of genetic variations between patients and normal individuals may provide a method of identifying the loci responsible for disease susceptibility (Hirschhorn and Daly 2005). One advantage of this method is that it does not need a large family. However, several limitations such as population structure, different levels of linkage disequilibrium (LD) (see below) in loci, and epistatic interaction of alleles may impose difficulties. Despite these limitations, there has been some success in identifying the association between polymorphisms and diseases (Table 1).

Unfortunately, however, this type of whole-genome approach to mapping requires the genotyping of thousands of samples. Although there are several high-throughput methods that are available for these studies, they are expensive, laborious and cannot be undertaken by many laboratories. Therefore, a different procedure called haplotype (collection of SNPs on a single chromosome at a locus that is inherited in blocks, Fig. 1B) analysis has been used to identify common disease genes (Hirschhorn and Daly 2005). Because the genome undergoes recombination involving large stretches of DNA, there may be several SNPs linked together in this large region of DNA. These closely linked SNPs may then be cotransmitted from generation to generation in these large blocks (Reich et al. 2001). This phenomenon is called LD. In this type of analysis, multiple genotypes are reduced to haplotypes and hence only a small number of SNPs are required to map the disease gene. Therefore, this method can be more effective in gene mapping and can also provide substantial statistical power in association studies (McVean et al. 2005). However, for conducting genetic association studies, putative polymorphism must be validated. The deleterious effect of SNPs should be evaluated in the context of a relevant

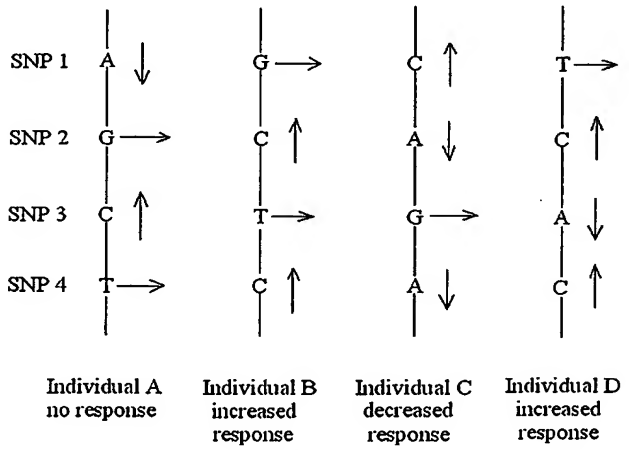


Fig. 2 A hypothetical haplotype and drug response. Individual *A* shows no response, individuals *B* and *D* show an increased response, whereas individual *C* shows a decreased response to a drug. The *horizontal arrows* denote no change in gene activity, whereas *upward* and *downward arrows* represent increased and decreased gene activity, respectively. Haplotype analysis may give a more accurate prediction of drug response

haplotype because it is more accurate than a single SNP (Fig. 2). Additionally, previous studies have shown that in the human genome many variants may be common to all populations and others may have a very restricted distribution (Salisbury et al. 2003). Hence its use in disease gene mapping requires additional research. In this regard, the recently characterized second type of DNA variation, called copy number polymorphisms (CNVs), which show marked variations among populations and individuals, may be helpful (Sharp et al. 2005; Locke et al. 2006).

Pharmacogenetics and pharmacogenomics

Single nucleotide polymorphism technologies are also applicable in the development of individualized medicine. Over the past 20 years it has become increasingly clear that genetic polymorphism in genes encoding drug-metabolizing enzymes, drug transporters and receptors contribute, at least in part, to the inter-individual variability in drug response (reviewed in Shastry 2003, 2004; Evans and Johnson 2001). These factors affect drug absorption, distribution, metabolism and excretion. As a result, some drugs work better in some patients than others, and some drugs may be highly toxic to certain patients (Ansari and Krajinovic 2007). This type of anti-drug reaction has been observed in several diseases, such as pulmonary hypertension, epilepsy, cardiac arrhythmia, renal cell carcinoma, leukemia and liver cancer (reviewed in Shastry 2006a, 2006b; Roses 2000). In order to understand the relationship between heritable changes in genes and

Table 2 A partial list of genetic polymorphisms associated with drug response

Gene	Name	Variant	Phenotypic effect
CYP 3A5	Cytochrome P-450	Splicing	Severely reduced activity
CYP 3A4	Cytochrome P-450	F189S	Reduced activity
LIP C	Hepatic lipase	C-514T	C/C genotype shows increased response to statin, variation in HDL-C levels
MTHFR	Methylene tetra-hydrofolate reductase	C677T	T/T genotype shows increased toxicity
HTR2A	Serotonin receptor 2A	H452Y	Reduced response to clozapine
ABCB1	p-Glycoprotein 1	C3435T	T/T patients may have less drug resistance
IL-10	Interleukin 10	A1082G	G/G individuals have better response to prednisone
GSTP1	Glutathione *S*-transferase	I105V	Increased survival for 5-fluorouracil
AGT	Angiotensinogen	M235T	Reduction of blood pressure
ADD1	Adducin 1	G460W	Increased response to diuretic in hypertensive patients
ABCG2	Efflux transporter of gefitinib	C421A	Gefitinib accumulates in heterozygotes
CDA	Cytidine deaminase	G208A	Severe drug toxicity to gemcitabine

inter-individual variations to drug response, two related fields namely pharmacogenetics and pharmacogenomics (Dervieux and Bala 2006) emerged and gained popularity in the late 1990s. They have undertaken massive studies on the genetic personalization of drug response (McLeod and Evans 2001). As a result, there are several high-density SNP maps of genes encoding proteins of medical importance (Iida et al. 2002, 2003), and now there is strong evidence (Table 2) that links SNPs to inter-individual differences in drug response (Nothen and Cichon 2002; Ansari and Krajinovic 2007). Patients with more active drug-metabolizing enzymes may require higher doses of the drug, and those who do not have an active enzyme may exhibit toxicity (Fig. 1C).

However, it should be noted that there are many negative results regarding the association of gene polymorphism and drug efficacy and toxicity. In addition to these negative results there are other problems. For example, drug metabolism or variation in drug response includes dozens of genes, and many of these often have multiple polymorphisms. Moreover, there are inducible genes, signaling molecules and environmental factors that may also contribute to variable drug response. The greatest challenge for the future (Roden et al. 2006) is to understand the genotypic–environmental factor interaction, ethnicity, inheritance patterns in drug response and how genetic variance responds to medicine. If the goals of pharmacogenetics and pharmacogenomics are fulfilled, it may allow clinicians to genetically subdivide and profile individual patients and treat each patient according to their genetic make-up. This type of medical practice may gradually replace the current trial-and-error-based selection of medicine in the future. However, at present, studies do not unambiguously prove the clinical value of pharmacogenetic testing.

Nutrigenetics and nutrigenomics

It is well known that certain monogenic disorders are associated with the interaction between variant genes and nutrients. The best example is phenylketonuria. Several recent population-based and intervention studies (Subbiah 2007; Ordovas and Mooser 2004; Ordovas et al. 2002a) support this gene–nutrient interaction. Polymorphism on its own may not have an effect, but nutrients may modulate the expression of genes. For instance, a significant variation in the low-density lipoprotein cholesterol level is shown to be associated with A-204C variant in CYP7 (Couture et al. 1999) gene, and high-density cholesterol concentration is determined by polymorphism C-514T in the hepatic lipase gene (Zhang et al. 2005; Couture et al. 2000). Similarly, the high-density lipoprotein cholesterol level is modulated by apolipoprotein A1 (APOA1) genetic polymorphism (-75G/A) in the promoter region (Ordovas et al. 2002b). Therefore, an understanding of the genetic make-up of an individual may lead to the development of an individualized diet. This may reduce diet-related disease risk more efficiently in some common multifactorial disorders (Ordovas and Mooser 2004). Because of this important relationship between gene–nutrient interactions and human health, two recently developed multidisciplinary fields, namely nutrigenetics and nutrigenomics, are exploring the possibility of developing personalized diets based on the genetic make-up of an individual.

SNPs in evolution

Genetic variants are not only considered to be responsible for disease risk and inter-individual differences, but also

molecular evolution. Genetic evolution in part depends upon a balance between natural selection and environmentally driven mutation. The natural selection will maintain and retain the amino acid type and position among species because these amino acids are critical for the protein function. Therefore, in a given set of homologous genes, certain amino acids are highly conserved, even among distantly related species that diverged hundreds of million years ago (Fig. 3). These conserved residues are evolving under strong selective pressure. Deleterious mutations that affect the biological functions of proteins are effectively eliminated by natural selection from the gene pool. The selection pressure against deleterious SNPs depends upon the molecular functions of proteins and those genes that encode transcriptional regulatory proteins are generally found to be under the strongest selective pressure (Ramensky et al. 2002).

Because SNPs are present at all levels of evolution, including the branch point of speciation, they can be used to study sequence variation among species. Additionally, the rate, type and site of substitution as well as the selection pressure on codons are not uniform throughout the given gene. Therefore, if genetic variants are fixed during evolution, then they may have either selection advantages for the organism, they may be neutral regarding the fitness, or they may be deleterious and thus cause pathology. Hence, a comparative genomic study of disease-associated SNPs can be used to understand the relationship between the pathology and evolution.

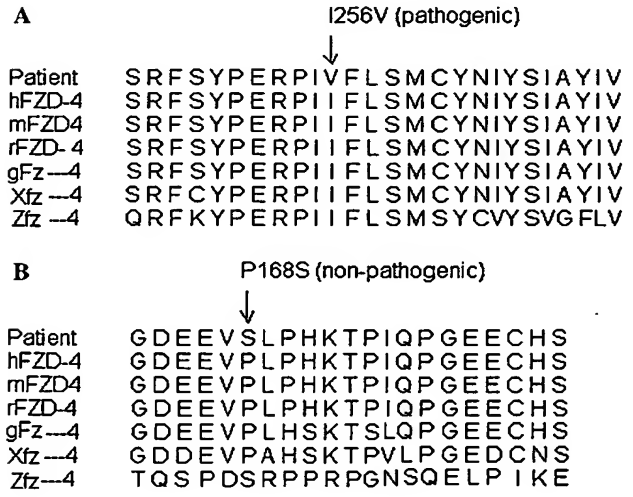


Fig. 3A–B Protein sequence alignment of the mutant part of the human frizzled-4 with that of other species. A conservative change in codon 256 causes pathology in the patient (**A**) whereas a radical change in a less conserved residue is nonpathogenic (**B**). *h*, human; *m*, mouse; *r*, rat; *X*, xenopus; *Z*, zebra fish; and *g*, chicken

Evolutionarily conserved regions more frequently contain disease-associated SNPs

The retention of variants by natural selection is considered to be an important step in evolution. According to the neutral theory of evolution (Kimura 1983), those amino acids that vary among species or those SNPs that do not occur in protein coding regions are either not subjected to natural selection or are under less selective pressure. This is because such amino acid changes can be tolerated and they only minimally affect protein functions. This may imply that such amino acids are more stable and less mutable. However, nSNPs in the coding regions of human genes may have phenotypic effects (Bao and Cui 2005) and may undergo natural selection. Hence, by comparing the rate ratio (omega) of nonsynonymous to synonymous changes (which is considered to be a measure of selective pressure on amino acid replacement mutations) in several proteins from several different species, raw evolutionary data can be generated.

In protein coding genes, patterns of selection can be inferred from amino acid substitution patterns (Jiang and Zhao 2006). One interesting example used to illustrate this is the patterns of distribution of disease-associated and nonpathogenic mutations in human genes. For instance, by using disease-associated mutation data and multiple species of phylogenetic lineage, it has been shown that disease-associated substitution (DAS) occurs more frequently in evolutionarily conserved positions (nonrandom distribution) than in positions that are undergoing variation (Subramanian and Kumar 2006; Miller and Kumar 2001). On the other hand, the opposite trend has been observed for silent and polymorphic mutations, and these are randomly distributed. These patterns are reinforcing the logic that the conserved region of the protein is under evolutionary pressure because these amino acids are critical to the proper functioning of a given gene. On the other hand, silent mutations have minimal affects on the organism because of their random distribution and may not be subjected to natural selection. However, polymorphic mutations of variable amino acids (nonconserved) may have moderate deleterious effects on the organism, and it is likely that such affects are tolerated and hence evolution may be more relaxed in these nucleotides. Similarly, a comparative study between human and chimpanzee genome indicates that some of the human specific traits could be due to positive selection, whereas loci for complex disorders could involve negative selection (Kehrer-Sawatzki and Cooper 2007; Patterson et al. 2006).

Another simple example is the myostatin gene (a negative regulator of skeletal muscle growth). When two different human populations and other mammals are compared for the myostatin gene (Saunders et al. 2006), the

number of highly conserved replacement mutations over the evolutionary time scale is greater (five) than the number of silent mutations (three). These data suggest a positive natural selection in the highly conserved region of the myostatin gene because, according to the neutral model of molecular evolution, the ratio of replacement to silent changes does not differ within and between species. However, at present it is not known what types of specific traits are associated with these five replacement changes and what kinds of selection advantages they may have for the species. Additional studies in the future may provide some answers to these questions.

Substitution patterns in the regulatory regions of DNA and noncoding RNA

Natural selection not only operates on protein coding genes but also at the RNA level and on the noncoding regions (regulatory regions) of the DNA. A similar distribution pattern to that discussed above has also been observed for DAS and SNPs in regulatory regions of genes and in RNAs that do not code for proteins (Keightley and Gaffney 2003). For example, it is estimated that at the genomic level, the deleterious point mutation rate is similar between noncoding and coding DNA. Moreover, deleterious mutations in noncoding DNA have quantitative effects, which means that these variations can produce complex genetic diseases (Keightley and Gaffney 2003). Similarly, using SNP genotypic data, it has been shown that negative selection in humans is stronger on conserved microRNA (miRNA binds to the target sites in the 3′-untranslated region of mRNA to repress the translation) binding sites than on other conserved sequence motifs in the 3′-untranslated regions (Saunders et al. 2007). This illustrates the importance of miRNAs to Darwinian fitness (Chen and Rajewsky 2006). Interestingly, a comparison between the miRNA and target sites shows a relatively low level of variation in the functional regions of miRNA and an appreciable level of variation in target sites. Some of these SNPs create novel target sites for miRNA and are found at relatively high frequencies in human populations. If some of these variants have functional effects, they may be involved in phenotypic differences and hence may undergo positive selection. Similarly, an evolutionary comparison using entire classes of mammalian sequences has provided other evidence for the relationship between the pathogenicity of RMRP (RNA component of the mitochondrial RNA processing ribonuclease) mutations and evolutionarily conserved sequences (Bonafe et al. 2005). Although this RNA does not code for a protein, some regions of the RNA are critical to protein binding. The encoding gene is remarkably conserved between species, but disease-causing mutations are once again found in highly conserved nucleotides whereas nonpathogenic variants are located in the nonconserved positions (evolution is more relaxed in these nucleotides). This is consistent with other examples discussed above for the protein coding genes.

Selection pressure is not uniform at amino acid sites

It should also be noted that there are differences in types of amino acid substitutions between species and diseases (Yang et al. 2000). For instance, among species glutamic acid is most commonly replaced by an aspartic acid (very similar) and phenylalanine is replaced by tyrosine. However, this trend has not been observed in disease. When a total of 4,236 mutations in 436 genes causing Mendelian disease (monogenic etiology) and 1,037 synonymous and nSNPs in 313 human genes are compared, a significantly larger contribution at arginine and glycine (also to some extent lysine) is observed in human genetic diseases (Vitkup et al. 2003). This is not the type of change accepted by natural selection. Additionally, a random mutation at tryptophan or cystein residues has the highest probability of causing a disease. This is in agreement with our understanding of their highest evolutionary contribution, which is nothing but their (trp and cys) involvement in determining the protein stability. Thus, selection pressure is not uniform among codons (Arbiza et al. 2006), and in many cases whenever a highly conserved codon is mutated it causes pathology (Fig. 3A).

Radical and less radical SNPs cause early- and late-onset diseases, respectively

Similar to the difference in types of amino acid substitutions between species and diseases, selection pressure also varies between species and disease depending on the properties of amino acids. Those amino acids that have larger chemical difference (radical) are more likely to produce disease phenotypes than those with smaller chemical properties (less radical). Amino acids that have smaller chemical properties are mostly observed among species. As mentioned above, it is the mutation with the larger chemical difference that is most likely to be removed from the population over a long period of time because they are likely to be deleterious. On the other hand, radical changes in variable positions (Fig. 3B) are more likely tolerated (they may not have large effects on protein functions) than in highly conserved positions. Hence these positions do not undergo strong selection. Interestingly, early-onset diseases (they are more damaging) are found to be associated with more radical amino acid mutations, and

as a consequence these positions are expected to undergo strong selection. In the same way, late-onset diseases are associated with less radical amino acid mutations and they are not abundant in evolutionarily conserved positions. These less radical amino acid mutations are often associated with common diseases such as diabetes and hypertension. Because they are involved in late-onset diseases, they may have smaller effect on fertility and hence these positions may not undergo strong natural selection. In short, comparative genomic studies between homologous gene sequences from both closely and distantly related species predict that evolution and DAS (pathology) are interrelated. Those residues that evolve under strong selective pressures are likely to be significantly associated with human disease (Arbiza et al. 2006). These types of studies also give us some understanding of the types of variations that can be tolerated in a given gene over time.

Substitution patterns and rates at the chromosome level

Although a lengthy discussion on this subject is not intended in this article, it is relevant to add that the evolutionary rates across the human chromosome are also not constant (Prendergast et al. 2007). Previous studies have predicted a relatively constant mutation rate across mammalian genomes. However, a recent analysis of human–mouse alignment suggests an approximately threefold difference in substitution rates across chromosomes. One of the factors that are found to be associated with mutation rates is the chromatin structure. The human genome contains two types of chromatin structures—closed and open. The open regions of genome are gene-dense and closed regions are relatively gene-poor (Gilbert et al. 2004). Housekeeping and tissue-specific genes are generally found in the more open and most closed regions of the genome, respectively. According to a recent study, the density of SNPs is higher in the most closed regions of the human genome, and genes in these regions also show the highest level of selection at synonymous sites. In fact, the average rate of nonsynonymous changes (d*N*) observed in human–mouse alignments is much higher in the most closed chromatin region of the genome than in the most open regions. Similarly, the ratio of nonsynonymous to synonymous substitution rates (d*N*/d*S*) is also higher, which indicates a strong selection. On the other hand, genes in the regions of open chromatin display the lowest mutation rates and the least constraints at the synonymous sites. However, the average synonymous rate (d*S*) for genes in relatively open chromatin is higher than that for genes in a closed chromatin structure. One of the explanations suggested by researchers for the lesser constraint in the regions of open chromatin is that open regions may be more accessible to repair mechanisms. On the other hand, as mentioned earlier, changes at synonymous sites do not affect the encoded amino acids. Therefore, a synonymous site would have to undergo relatively strong selection to evolve in a non-neutral condition. It is also possible that synonymous sites may experience constraints because they may have a role in RNA stability or splicing.

Fitness, gene pool and functional redundancy

These types of SNPs studies (comparison of relative fixation rates of silent and nSNPs) may allow us to trace the branch point of an evolutionary tree. At this branch point, the variants must have become advantageous for the species and fixed in the gene pool (Zhang et al. 2006). According to comparative genomics, those sequences that contribute to the fitness of an organism evolve slowly. For example, selenoproteins play an important role in antioxidant defense. When polymorphisms of six genes, namely glutathione peroxidases (GPX1, GPX2, GPX3, GPX4), thioredoxine reductase 1 (TXNRD1) and selenoprotein P (SEPP1), were compared in 102 individual populations representing four major ethnic groups, evidence for positive selection was found at the GPX1 locus (Foster et al. 2006). However, in the remaining five genes there was no strong evidence for selection and hence they must have adopted the neutral equilibrium model of evolution. This may imply that they are functionally redundant. It is not clear at present whether this selective pressure on GPX1 is exerted to protect the genome from damaging oxidants or to reduce susceptibility to oxidative stress in erythrocytes, where it is mostly expressed, or both.

Similarly, the ability to digest lactose (present in milk) usually disappears in childhood in most human populations. However, in European-derived populations, lactase activity persists into adulthood. This type of lactase persistence could be due to multiple causes and it may also depend on the population under study. One interesting finding, however, is that, when a region of 3.2 Mbp around the lactase gene consisting of 101 SNPs were typed in northern European and African populations, two alleles were found to be tightly associated with lactase persistence (Trishkoff et al. 2007; Coelho et al. 2005; Bersaglieri et al. 2004). This association could be due to a strong positive selection because of animal domestication and adult milk consumption (advantages to the organism), and hence it is fixed in the gene pool. In contrast, the human mannose binding lectin (MBL-2) allele (a member of the collectin protein family that binds a broad range of microorganisms) occurs at a high frequency worldwide (Verdu et al. 2006). This allele produces little or no protein and was shown to result from human migration and genetic drifts. This

evolutionary neutrality (with respect to fitness) of MBL-2 may also suggest that the MBL-2 allele is functionally redundant in the host human defense.

Additional factors that may also contribute to the evolution of the human genome may include DNA methylation, genome duplication, deletions, insertions and the presence of introns (Tang et al. 2006). Insertions and deletions are collectively known as indels and they are approximately 300 bp in length. Because of their high frequency and wide distribution, indels are considered to be the strong driving force of evolution. In addition, the distribution patterns of DAS and nSNPs also show that positions that have many indels in other species contain more nSNPs than DAS. This is not due to the mutation rate, because an excess of nSNPs would be expected in positions with many indels if it was, and that is not found to be the case. Future studies using a recently characterized second type of DNA variation, called CNVs, which show marked variations among populations and individuals, may be helpful (Sharp et al. 2005; Locke et al. 2006) in understanding genetic diversity and evolution.

Concluding remarks

After the *First International Meeting on SNPs* in 1998, it was realized that SNP technologies may have an impact on healthcare. There is no doubt that clinicians, geneticists, patients and the public will benefit from the identification of genes underlying polygenic diseases and adverse drug reactions. Over the past ten years, tremendous progress has been made in cataloging human sequence variations since this high-density map will provide the necessary tools to develop genetically based diagnostic and therapeutic tests. When more functional polymorphisms have been identified, it may be possible to develop useful genetic markers as well as personalized medicines. If the concept of individualized medicine becomes more realistic, every newborn child in the neonatal unit may be genotyped in a routine procedure (similar to a blood transfusion procedure) for improved treatment. The newly developed fields of toxicogenomics, pharmacogenetics and nutrigenetics are rapidly advancing to achieve their goals.

Another interesting aspect of SNPs is that they can also be used to understand the molecular mechanisms of sequence evolution. Natural selection will maintain the amino acid type and retain the amino acid position among species because these amino acids are critical to protein function. Deleterious mutations that affect the biological functions of proteins are effectively being eliminated by natural selection from the gene pool. As discussed above, there is a clear evolutionary relationship between the positions and types of neutral and DAS in the human genome. Residues that evolve under strong selective pressure are found to be significantly associated with human diseases. These patterns are clearly different among species. In short, nucleotide substitutions that are fixed during evolution are either in some way advantageous for the organism, remain neutral regarding fitness, or become deleterious and thus cause pathology. Therefore, evolution and disease-causing nucleotide substitutions can be considered to be related to one another. In the future, it is hoped that research will uncover methods of making SNP markers useful tags for medical testing. Finding out how SNPs affect the health of an individual and then transforming this knowledge into the development of new medicines will undoubtedly revolutionize the treatments of the most common devastating disorders. At the same time, this knowledge will also help us to uncover the secrets of human genome evolution.

Acknowledgments My apologies to those whose work or original publications could not be cited in this short article because of space limitations.

References

Ansari M, Krajinovic M (2007) Pharmacogenomics in cancer treatment defining genetic bases for inter-individual differences in responses to chemotherapy. Curr Opin Pediatr 19:15–22

Arbiza L, Duchi S, Montaner D, Burguet J, Pantoga-Uceda D, Pineda-Lucena A, Dopazo J, Dopazo H (2006) Selective pressures at a codon level predict deleterious mutations in human disease genes. J Mol Biol 358:1390–1404

Bao L, Cui Y (2005) Prediction of the phenotypic effects of non-synonymous single nucleotide polymorphisms using structural and evolutionary information. Bioinformatics 21:2185–2190

Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, Rhodes M, Reich DE, Hirschhorn JN (2004) Genetic signatures of strong recent positive selection at the lactase gene. Am J Hum Genet 74:1111–1120

Bonafe L, Dermitzakis ET, Unger S, Greenberg CR, Campos-Xavier BA, Zankl A, Ucla C, Antonarakis SE, Superti-Furga A, Reymond A (2005) Evolutionary comparison provides evidence for pathogenicity of RMRP mutations. PLoS Genet 1:444–454

Chen K, Rajewsky N (2006) Natural selection on human microRNA binding sites inferred from SNP data. Nat Genet 38:1452–1456

Coelho M, Luiselli D, Bertorelle G, Lopes AT, Seixas S, Destro-Bisol G, Rocha J (2005) Microsatellite variation and evolution of human lactase persistence. Hum Genet 117:329–339

Couture P, Otvos JD, Cupples LA, Wilson PW, Schaefer EJ, Ordovas JM (1999) Association of the A-204C polymorphism in the cholesterol 7 alpha-hydroxylase gene with variations in plasma low-density lipoprotein cholesterol levels in the Framingham offspring. J Lipid Res 40:1883–1889

Couture P, Otvos JD, Cupples LA, Lahoz C, Wilson PW, Schaefer EJ, Ordovas JM (2000) Association of the C-514T polymorphism in the hepatic lipase gene with variations in lipoprotein subclass profiles: The Framingham Offspring Study. Arterioscler Thromb Vasc Biol 20:815–822

Dervieux T, Bala MV (2006) Overview of the pharmacoeconomics of pharmacogenetics. Pharmacogenomics 7:1175–1184

Emilien G, Ponchon M, Caldas C, Isacson O, Maloteaux J-M (2000) Impact of genomics on drug discovery and clinical medicine. Q J Med 93:391–423

Evans WE, Johnson JA (2001) Pharmacogenomics: the inherited basis for interindividual differences in drug response. Annu Rev Genomics Hum Genet 2:9–39

Foster CB, Aswath K, Chanock SJ, McKay HF, Peters U (2006) Polymorphism analysis of six selenoprotein genes: support for a selective sweep at the glutathione peroxidase 1 locus (3p21) in Asian populations. BMC Genet 7:56–63

Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP, Bickmore WA (2004) Chromatin architecture of the human genome: gene rich domains are enriched in open chromatin fibers. Cell 118:555–566

Gray IC, Campbell DA, Spurr NK (2000) Single nucleotide polymorphism as tools in human genetics. Hum Mol Genet 9:2403–2408

Halushka MK, Fan JB, Bentley K, Hsie L, Shen NP, Weder A, Cooper R, Lipshutz R, Chakravarti A (1999) Patterns of single nucleotide polymorphisms in candidate genes for blood pressure homeostasis. Nat Genet 22:239–247

Hirschhorn JN, Daly MJ (2005) Genome-wide association studies for common diseases and complex traits. Nat Rev Genet 6:95–108

Iida A, Saito S, Sekine A, Mishima C, Kitamura Y, Kondo K, Harigae S, Osawa S, Nakamura Y (2002) Catalog of 605 single nucleotide polymorphisms (SNPs) among 13 genes encoding human ATP binding cassettes transporters: ABCA4, ABCA7, ABCA8, ABCD1, ABCD2, ABCD4, ABCE1, ABCF1, ABCG1, ABCG2, ABCG4, ABCG5 and ABCG8. J Hum Genet 47:285–310

Iida A, Saito S, Sekine A, Mishima C, Kitamura Y, Kondo K, Harigae S, Osawa S, Nakamura Y (2003) Catalog of 668 SNPs detected among 31 genes encoding potential drug targets on the cell surface. J Hum Genet 48:23–46

Jiang C, Zhao Z (2006) Mutational spectrum in the recent human genome inferred by single nucleotide polymorphisms. Genomics 88:527–534

Kehrer-Sawatzki H, Cooper DN (2007) Understanding the recent evolution of the human genome: insights from human–chimpanzee genome comparisons. Hum Mutat 28:99–130

Keightley PD, Gaffney DJ (2003) Functional constraints and frequency of deleterious mutations in non-coding DNA of rodents. Proc Natl Acad Sci USA 100:13402–13406

Kimchi-Sarfaty C, Oh JM, Kim I-W, Kim IW, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM (2007) A silent polymorphism in the MDR1 gene changes substrate specificity. Science 315:525–528

Kimura M (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge, UK

Locke DP, Sharp AJ, McCarroll SA, McGrath SD, Newman TL, Cheng Z, Schwartz S, Albertson DG, Pinkel D, Altshuler DM, Eichler EE (2006) Linkage disequilibrium and heritability of copy-number polymorphisms within duplicated regions of the human genome. Am J Hum Genet 79:275–290

Martin N, Boomsma D, Machin G (1997) A twin pronged attacks on complex traits. Nat Genet 17:387–392

McLeod HL, Evans WE (2001) Pharmacogenomics: unlocking the human genome for better drug therapy. Annu Rev Pharmacol Toxicol 41:101–121

McVean G, Spencer CCA, Chaix R (2005) Perspectives on human genetic variation from the HapMap Project. PLoS Genet 1:413–418

Miller MP, Kumar S (2001) Understanding human disease mutations through the use of interspecific genetic variation. Hum Mol Genet 10:2319–2328

Nothen MM, Cichon S (2002) Linking single nucleotide polymorphisms. Pharmacogenetics 12:89–90

Ordovas JM, Mooser V (2004) Nutrigenomics and nutrigenetics. Curr Opin Lipidol 15:101–108

Ordovas JM, Corella D, Demissie S, Cupples LA, Couture P, Coltell O, Wilson PW, Schaefer EJ, Tucker KL (2002a) Dietary fat intake determines the effect of a common polymorphism in the hepatic lipase gene promoter on high-density lipoprotein metabolism: evidence of a strong dose effect in this gene–nutrient interaction in the Framingham study. Circulation 106:2315–2321

Ordovas JM, Corella D, Cupples LA, Demissie S, Kelleher A, Coltell O, Wilson PW, Schaefer EJ, Tucker K (2002b) Polyunsaturated fatty acids modulates the effects of the APOA1G-A polymorphism on HDL-cholesterol concentrations in a sex-specific manner: Framingham study. Am J Clin Nutr 75:38–46

Patterson N, Richter DJ, Gnerre S, Lander ES, Reich D (2006) Genetic evidence for complex speciation of humans and chimpanzees. Nature 441:1103–1108

Prendergast JGD, Campbell H, Gilbert N, Dunlop MG, Blackmore WA, Semple CAM (2007) Chromatin structure and evolution in the human genome. BMC Evol Biol 7:72–84

Ramensky V, Bork P, Sunyaev S (2002) Human non-synonymous SNPs: server and survey. Nucleic Acids Res 30:3894–3900

Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, Richter DJ, Lavery T, Kouyoumijian R, Farhadian SF, Ward R, Landers ES (2001) Linkage disequilibrium in the human genome. Nature 411:199–204

Roden DM, Altman RB, Benowitz NL, Flockhart DA, Giacomini KM, Johnson JA, Krauss RM, McLeod HL, Ratain MJU, Relling MV, Ring HZ, Shuldiner AR, Weinshilboum RM, Weiss ST (2006) Pharmacogenomics: challenges and opportunities. Ann Intern Med 145:749–757

Roses AD (2000) Pharmacogenetics and the practice of medicine. Nature 405:857–865

Salisbury BA, Pungliya M, Choi JY, Jiang RH, Sun XJ, Stephens JC (2003) SNP and haplotype variation in the human genome. Mutat Res-Fund Mol Mech Mutagenesis 526:53–61

Saunders MA, Good JM, Lawrence EC, Ferrell RE, Li W-H, Nachman MW (2006) Human adaptive evolution of myostatin (GDF8), a regulator of muscle growth. Am J Hum Genet 79:1089–1097

Saunders MA, Liang H, Li WH (2007) Human polymorphism at microRNAs and microRNA target sites. Proc Natl Acad Sci USA 104:3300–3305

Sharp AJ, Locke DP, McGrath SD, Cheng Z, Bailey JA, Vallente RU, Pertz LM, Clark RA, Schwartz S, Segraves R, Oseroff VV, Albertson DG, Pinkel D, Eichler EE (2005) Segmental duplications and copy-number variation in the human genome. Am J Hum Genet 77:78–88

Shastry BS (2002) SNP alleles in human disease and evolution. J Hum Genet 47:561–566

Shastry BS (2003) SNPs and haplotypes: genetics markers for disease and drug response. Int J Mol Med 11:379–382

Shastry BS (2004) Role of SNP/haplotype map in gene discovery and drug development: an overview. Drug Dev Res 62:142–150

Shastry BS (2005) Genetic diversity and new therapeutic concepts. J Hum Genet 50:321–328

Shastry BS (2006a) Pharmacogenetics and the concept individualized medicine. Pharmacogenomics J 6:16–21

Shastry BS (2006b) Role of SNPs and haplotypes in human disease and drug development. In: Ozkan M, Heller MJ, Ferrari M (eds) Micro/nano technology in genomics and proteomics, vol II. Springer, New York, pp 447–458

Subbiah MT (2007) Nutrigenetics and nutriceuticals: the next wave riding on personalized medicine. Transl Res 149:55–61

Subramanian S, Kumar S (2006) Evolutionary anatomies of position and types of disease associated and neutral amino acid mutations in the human genome. BMC Genomics 7:306–312

Tang CS, Zhao YZ, Smith DK, Epstein RJ (2006) Intron length and accelerated 3′ gene evolution. Genomics 88:682–689

Trishkoff SA, Reed FA, Ranciaro A, Voight BF, Babbitt CC, Silverman JS, Powell K, Mortensen HM, Hirbo JB, Osman M, Ibrahim M, Omar SA, Lema G, Nyambo TB, Ghori J, Bumpstead S, Pritchard JK, Wray GA, Deloukas P (2007) Convergent adaptation of human lactase persistent in Africa and Europe. Nat Genet 39:31–40

Verdu P, Barreiro LB, Patin E, Gessain A, Cassar O, Kidd JR, Kidd KK, Behar DM, Froment A, Heyer E, Sica L, Casanova JL, Abel L, Quintana-Murci L (2006) Evolutionary insights into the high worldwide prevalence of MBL2 deficiency alleles. Hum Mol Genet 15:2650–2658

Vitkup D, Sander C, Church GM (2003) The amino acid mutational spectrum of human genetic disease. Genome Biol 4: R72-R80

Yang Z, Nielsen R, Goldman N, Pedersen AM (2000) Codon substitution models for heterogeneous selection pressure at amino acid sites. Genomics 155:431–449

Zhang C, Bailey DK, Awad T, Liu G, Xing G, Cao M, Valmeekam V, Retief J, Matsuzaki H, Taub M, Seielstad M, Kennedy GC (2006) A whole genome long-range haplotype (WGLRH) test for detecting imprints of positive selection in human populations. Bioinformatics 22:2122–2128

Zhang C, Lopez-Ridaura R, Rimm EB, Rifai N, Hunter DJ, Hu FB (2005) Interaction between the -514 C to T polymorphism of the hepatic lipase gene and life-style factors in relation to HDL concentrations among US diabetic men. Am J Clin Nutr 81:1429–1435